

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 1999, 18:22:40 ; Search time 556.3 Seconds  
(without alignments)  
12036.795 Million cell updates/sec

Title: US-08-699-716a-1

Perfect score: 1566  
Sequence: 1 ATGGGCATCATCATCATCA.....ATGACACGCTCGTAAATGA 1566

Scoring table: IDENTITY\_NUC

Searched: 780561 seqs, 2137953050 residues

Database :

GenEmbl:\*  
1: gb\_ba1:\*  
2: gb\_ba2:\*  
3: gb\_om:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_ov:\*  
7: gb\_pl1:\*  
8: gb\_pl2:\*  
9: gb\_pl1:\*  
10: gb\_pl2:\*  
11: gb\_pl3:\*  
12: gb\_pl3:\*  
13: gb\_pl3:\*  
14: gb\_sy:\*  
15: gb\_sy:\*  
16: gb\_un:\*  
17: gb\_un:\*  
18: em\_fun:\*  
19: em\_hlg:\*  
20: em\_hum1:\*  
21: em\_hum2:\*  
22: em\_hum3:\*  
23: em\_hum4:\*  
24: em\_hum4:\*  
25: em\_hum4:\*  
26: em\_hum4:\*  
27: em\_hum4:\*  
28: em\_hum4:\*  
29: em\_hum4:\*  
30: em\_hum4:\*  
31: em\_hum4:\*  
32: em\_hum4:\*  
33: gb\_hlg1:\*  
34: gb\_hlg2:\*  
35: gb\_hlg3:\*  
36: gb\_hlg4:\*  
37: em\_hum4:\*  
38: em\_hum4:\*  
39: em\_hum4:\*  
40: em\_hum4:\*  
41: gb\_hlg3:\*  
42: gb\_hlg3:\*  
43: gb\_hlg4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
-----						

## BEST AVAILABLE COPY

1	1473.6	94.1	1530	5	A56814	A56814 Sequence 22
2	1410.6	90.1	1462	5	A56808	A56808 Sequence 16
3	979.4	62.5	2100	1	YEPICR	M26405 Yersinia pe
4	979.4	62.5	70305	1	YPCD1	AL117189 Yersinia
5	979.4	62.5	70504	2	AF053946	AF053946 Yersinia
6	979.4	62.5	70559	2	AF074612	AF074612 Yersinia
7	976.8	62.4	1014	5	A46411	A46411 Sequence 1
8	976.8	62.4	1014	5	A56793	A56793 Sequence 1
9	974.8	62.2	1014	5	A46413	A46413 Sequence 3
10	974.8	62.2	1014	5	A56795	A56795 Sequence 3
11	957	61.1	981	1	YEPICR	X56802 Y.pseudotub
12	953.8	60.9	2201	1	YEPICR	M57893 Yersinia ps
13	927.4	59.2	63673	2	AF102990	AF102990 Yersinia
14	916.2	58.5	975	1	YEPICR	X56801 Y.enterecol
15	914.6	58.4	975	1	YEPICR	X56797 Y.enterecol
16	913	58.3	975	1	YEPICR	AL117211 Yersinia
17	911.4	58.2	975	1	YEPICR	AF053947 Yersinia
18	890.4	56.9	1002	1	YEPICR	AF074611 Yersinia
19	882.4	56.3	1002	1	YEPICR	A56812 Sequence 20
20	516.8	33.0	5383	1	YPCAF	S38727 ICRGVH oper
21	516.8	33.0	96210	2	AF010149	AF010149 Pseudomon
22	516.8	33.0	100984	2	AF074611	AF074611 Yersinia
23	516.8	33.0	100990	2	AF074611	AF074611 Yersinia
24	516.8	33.0	547	5	A56812	A56812 Sequence 20
25	217.6	13.9	240	1	S38727	S38727 ICRGVH oper
26	74	4.8	4168	2	AF010149	AF010149 Pseudomon
27	74	4.7	660	14	AF081364	AF081364 Synthetic
28	53.4	3.4	1402	5	AR000030	AR000030 Sequence
29	53.4	3.4	521	11	S75264	S75264 WT1-Wilms'
30	48.2	3.1	85779	7	SE011856	AJ011856 Saccharom
31	48	3.1	7218	5	I66494	I66494 Sequence 14
32	47.8	3.1	256172	34	AC005139	AC005139 Plasmodu
33	47.8	3.1	155573	34	AC007926	AC007926 Trypanoso
34	47.8	3.1	110000	42	AC005140	AC005140 Plasmodu
35	46.8	3.0	293431	33	PFMAL13P4	AL049181 Plasmodu
36	46.8	3.0	110000	33	PFMAL4P1_0	AL034457 Plasmodu
37	46.4	3.0	1158	35	AF044866	AF044866 Phoeblis s
38	45.4	2.9	314303	42	AC006279	AC006279 Plasmodu
39	45.2	2.9	21202	21	PFSC04088	AL101073 Plasmodu
40	45.2	2.9	224448	33	PFMAL4P4	AL035477 Plasmodu
41	44.8	2.9	1978	35	DDIDPKA	M59744 D. discoiden
42	44.8	2.9	12029	36	AE001426	AE001426 Plasmodu
43	44.4	2.8	321003	33	PFMAL4P3	AL035476 Plasmodu
44	44.2	2.8	211831	42	AC005308	AC005308 Plasmodu
45	44.2	2.8	190112	43	AC011454	AC011454 Homo sapi

### ALIGNMENTS

RESULT 1	A56814	Sequence 22 from Patent WO9628551.	DNA	PAT	03-MAR-1998
LOCUS	A56814	1530 bp			
DEFINITION	A56814				
ACCESSION	93712827				
NID	93712827				
VERSION	A56814.1	GI:3712827			
KEYWORDS	Yersinia pestis.				
SOURCE	Yersinia pestis.				
ORGANISM	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Yersinia.				
REFERENCE	1 (bases 1 to 1530)				
AUTHORS	Titball,R.W., Williamson,E.D., Leary, Sophie,E.C., Oyston,P.C., Bennett and Alice,M.				
TITLE	VACCINES FOR PLAGUE				
JOURNAL	Patent: WO 9628551-A 22 19-SEP-1996;				
COMMENT	SECR DEFENCE (GB)				
FEATURES	Other publication ZA 9602036 960716				
	Location/Qualifiers				
	1..1530				
	/organism="Yersinia pestis"				





REFERENCE 1 (bases 1 to 2100)  
AUTHORS Price,S.B., Leung,K.Y., Barve,S.S. and Straley,S.C.  
TITLE Molecular analysis of *lcrGVH*, the *v* antigen operon of *Yersinia pestis*  
JOURNAL J. Bacteriol. 171, 5646-5653 (1989)  
MEDLINE 90008806  
COMMENT Computer readable copy of sequence [J. Bacteriol. (1989) In press]  
07-AUG-1989.

FEATURES  
source location/Qualifiers  
1..2100  
/organism="Yersinia pestis"  
/strain="KIM5"  
/db\_xref="taxon:632"  
-35\_signal 147..152  
-10\_signal 174..179  
gene 192..479  
/gene="lcrG"  
CDS 192..479  
/gene="lcrG"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="AA027640.1"  
/db\_xref="PID:g155449"  
/translation="MKSSHPEDYDKTLKQELAIAADSHRAKLLQEMCADIGLTPKAV  
MKIFAGRSAGEIKPAERELDEIKRQREPOPHYDGRKRPKPTMMRGQIT"  
463..466  
/gene="lcrG"  
463..1461  
/gene="lcrV"  
481..1461  
/gene="lcrV"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="AA027641.1"  
/db\_xref="PID:g155450"  
/db\_xref="GI:153450"  
/translation="MIRAYDNPQHFIEDLEKRVEDLTGHSSVLELVOLVYKKNID  
DISIKYPRKDESEVAVITDIELKILAYLPEDAILKGHYDNONGIKRYK  
EFLSSPNTOEILRAEVAHMFSLADRIDDLIKYVDSNMHHGDARSKLRELIAT  
TAEIKYVIOAEIKNLKSSGTTNIDKRSINMDKLYCTDEITKAAAEKILSEK  
MPOTTIOVDSSEKIKYIKDFLGSSEKRTALGVLKNSIYKNDNNLSFPAITTCSDK  
SRPLNDVSOQTQLSDITSRFNSAIEALNRFLOKXDSVMORLLDRTSGK"  
1465..1468  
/gene="lcrH"  
1465..1980  
/gene="lcrH"  
1474..1980  
/gene="lcrH"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="AA027642.1"  
/db\_xref="PID:g155451"  
/db\_xref="GI:153451"  
/translation="MOEETDIOEYOLAMESFLKGGGTIAMLEISSDTLEOYSLAF  
NOYOSGYEDAHKVFQALCYVDHSRFFGLGACRQAMQYDLAIHSYSGAIMDK  
EPREFPAAECLLOKGLAEASGLFLAQLIADKTEFKELSTRVSMLEAIKLKMK  
EHCEVDNP"

BASE COUNT 688 a 423 c 461 g 528 t  
ORIGIN

Query Match 62.5% Score 979.4; DB 1; Length 2100;  
Best Local Similarity 99.9%; Pred. No. 3.7e-204;  
Matches 980; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 586 ATGATTAGACCTAGCAACAACCCACACATTTTATTTGAGATCTGAGAAAAGTTAGG 645  
|||||  
Db 481 ATGATTAGACCTAGCAACAACCCACACATTTTATTTGAGATCTGAGAAAAGTTAGG 540

QY 646 GTGGAACAACCTACTGCTCATGCTTCTTCAGCTTTAGAGAATTGGTTCAGTTAGTCAAA 705

|||||  
Db 541 GTGGAACAACCTACTGCTCATGCTTCTTCAGCTTTAGAGAATTGGTTCAGTTAGTCAAA 600  
QY 706 GATAAAAATATAGATATTTTCATTAATATGATCCGAAAAGATTCGAGGTTTGGC 765  
|||||  
Db 601 GATAAAAATATAGATATTTTCATTAATATGATCCGAAAAGATTCGAGGTTTGGC 660  
QY 766 AATGAGTAATTTACTGATGATATGGAATTGCTCAAGAAAATCCTAGTTATTTCTACC 825  
Db 661 AATGAGTAATTTACTGATGATGATGGAATTGCTCAAGAAAATCCTAGTTATTTCTACC 720  
QY 826 GAGATACCATTTCTTAAGGGGGGTCATTATGACAAACCACTGCAGAAATGGCATCAACGCA 885  
Db 721 GAGATACCATTTCTTAAGGGGGGTCATTATGACAAACCACTGCAGAAATGGCATCAACGCA 780  
QY 886 GTAAAGAGTTCCTTGAATCATCCGCCAATACACAAATGGGAATTCGGCGGTTTCATGCA 945  
Db 781 GTAAAGAGTTCCTTGAATCATCCGCCAATACACAAATGGGAATTCGGCGGTTTCATGCA 840  
QY 946 GTAATGCATTTCTTTAACCGCGGATCGATGATGATGATTTTGAAGTGATTTGTT 1005  
Db 841 GTAATGCATTTCTTTAACCGCGGATCGATGATGATGATTTTGAAGTGATTTGTT 900  
QY 1006 GATTCATGATTCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1065  
Db 901 GATTCATGATTCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
QY 1066 ACCCGCAATTAAGATTTTATTCAGTTATTCAGCCCAATTAAGCACTCTGTAGT 1125  
Db 961 ACCCGCAATTAAGATTTTATTCAGTTATTCAGCCCAATTAAGCACTCTGTAGT 1020  
QY 1126 AGTGGACCATTAATATTCATGATTAATTCATTAATTCATGATTAATTAATGAT 1185  
Db 1021 AGTGGACCATTAATATTCATGATTAATTCATTAATTCATGATTAATTAATGAT 1080  
QY 1186 TATACAGATGAAGATTTTAAAGCCAGGAGAGTACCAAAATTCGAGAAAATGCGCT 1245  
Db 1081 TATACAGATGAAGATTTTAAAGCCAGGAGAGTACCAAAATTCGAGAAAATGCGCT 1140  
QY 1246 CAACACCATTCAGTGGATGGAGCGAGCAAAAATATGCTTCGATTAAGCACTTTCTT 1305  
Db 1141 CAACACCATTCAGTGGATGGAGCGAGCAAAAATATGCTTCGATTAAGCACTTTCTT 1200  
QY 1306 GGAAGTAGAATAAAGAACCGGGGCGTGGTATCTGAAAACTCATCTCTTAAT 1365  
Db 1201 GGAAGTAGAATAAAGAACCGGGGCGTGGTATCTGAAAACTCATCTCTTAAT 1260  
QY 1366 AAGATATAATGATATATATCTCATCTTGGCCACCACTGCTCGGATAGTCCAGGCGGCTC 1425  
Db 1261 AAGATATAATGATATATATCTCATCTTGGCCACCACTGCTCGGATAGTCCAGGCGGCTC 1320  
QY 1426 AACGACTTGGTTAGCCAAAACAACTACGCTGCTGATTTACATACAGCTTTAATTTCA 1485  
Db 1321 AACGACTTGGTTAGCCAAAACAACTACGCTGCTGATTTACATACAGCTTTAATTTCA 1380  
QY 1486 GCTATTAAGCACTGAACCGTTTCATTCAGAAATATGATTCAGTGAACGCTGCTCA 1545  
Db 1381 GCTATTAAGCACTGAACCGTTTCATTCAGAAATATGATTCAGTGAACGCTGCTCA 1440  
QY 1546 GATGACACGTCGTGTAATGA 1566  
Db 1441 GATGACACGTCGTGTAATGA 1461

RESULT 4  
YPCD1/c YPCD1 70305 bp DNA  
LOCUS Yersinia pestis plasmid pCD1.  
DEFINITION  
ACCESSION AL117189  
NID 95832423  
VERSION AL117189.1 GI:5832423  
KEYWORDS  
Chaperone; cytotoxic effector; IS100; IS1616; IS1617; lcr;  
low-calcium response; syc; targeted effector; transposase; type III

secretion; V antigen; virulence; y1p; yop; ysc.  
Yersinia pestis.  
Yersinia pestis  
Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;  
Yersinia.  
1 (bases 1 to 70305)  
Karlyshev,A.V. and Wren,B.W.  
Unpublished  
2 (bases 1 to 70305)  
Baker,S.G. and Mungall,K.  
Unpublished  
3 (bases 1 to 70305)  
James,K.D., Parkhill,J., Barrell,B.G. and Randalream,M.A.  
Direct Submission  
Submitted (03-SEP-1999) Yersinia pestis sequencing project, Sanger  
Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA  
E-mail: barrell@sanger.ac.uk DNA supplied by Dr. Andrey Karlyshev  
and Prof. Brendan Wren, [3]. Department of Infectious and Tropical  
Diseases, London School Of Hygiene and Tropical Medicine, Keppel  
Street, London WC1E 7HT  
Notes:  
Yersinia pestis sequencing at The Sanger Centre is funded by  
Beowulf Genomics  
Details of Y. pestis sequencing at the Sanger Centre are available  
on the World Wide Web.  
(URL: [http://www.sanger.ac.uk/Projects/Y\\_pestis/](http://www.sanger.ac.uk/Projects/Y_pestis/))  
CDS are numbered using the following system eg YPCP1.01c. yp (Y.  
pestis), PCP1 (plasmid name), .01 (first CDS), c (complementary  
strand).  
The more significant matches with motifs in the PROSITE database  
are also included but some of these may be fortuitous.  
The length in codons is given for each CDS.  
Usually the highest scoring similarity to other CDS in the database.  
CDS which show significant similarity to other CDS are given  
where the position of possible ribosome binding site sequences are given  
CAUTION: We may not have predicted the correct initiation codon.  
Where possible we choose an initiation codon (atg, gtg, ttg or  
(atc) which is preceded by an upstream ribosome binding site  
sequence (optimally 5-13bp before the initiation codon). If this  
cannot be identified we choose the most upstream initiation codon.  
Location/Qualifiers  
1. .70305  
/organism="Yersinia pestis"  
/plasmid="PCD1"  
/strain="CO-92 Biovar Orientalis"  
/db\_xref="taxon:632"  
1. .1956  
/note="IS100 element"  
88. .1110  
/gene="YPCD1.01"  
88. .1110  
/codon\_start=1  
/label="table=11"  
/label="YPCD1.01"  
/note="YPCD1.01, probable transposase, len: 340 aa;  
putative insertion sequence IS100, identical to  
corresponding CDS from Y.pestis KIMS PCD1 (EMBL:AF053946,  
AF074612) (340 aa). fasta scores: opt: 2328 z-score:  
3808.9 E(): 0, 100.0% identity in 340 aa overlap. Similar  
to many others e.g. TRA0\_ECOLI (EMBL:X14793), istA, E.coli  
transposase for insertion sequence element IS21 (390 aa)  
(33.1% identity in 329 aa overlap). Contains pfam match to  
entry PF00239 recombinase, site-specific recombinases,  
score 25.70, E-value 4.8e-06. Contains probable  
helix-turn-helix motif at aa 19-40 (Score 2045, +6.15 SD)"  
/codon\_start=1  
/transl\_table=11  
/label="YPCD1.01"  
/product="putative transposase"  
/protein\_id="CAB54878.1"  
/db\_xref="pid:e1549591"  
/db\_xref="pid:g5832424"  
/db\_xref="gi:5832424"  
/translation="MVTFTVMEIKILHRQGSRAIARELGISRTVARYIQAQSEPP

PKYTPRPAVASLIDEDYIRQRIADAHPIKIPATVIAREIRDQYRGCMITLAEIR  
YSLVQEOEPNAPFETEPGRQMVDMGMKRNRSPLHFAVLGYSMLYIEFTNMR  
XLTLETCNHNAPREFFGVREVLVYNNMTVYQRAYOGAFGRFSPSLMOFGKENGFS  
PLICRPNQAFQTKGVKREVMYQNSFYIPLMRLLDMGTVDFVTANRGLIWLIDVA  
NORKHETIOARPCDWMLEQOSMLALPPEKKEDVHLDENLVNFKHPLHRLSTYDS  
FCRGVA"  
112. .195  
/gene="YPCD1.01"  
/note="Pfam match to entry PF00239 recombinase,  
site-specific recombinases, score 25.70, E-value 4.8e-06"  
1110. .1889  
/gene="YPCD1.02"  
1110. .1889  
/gene="YPCD1.02"  
/note="YPCD1.02, probable transposase, len: 259 aa;  
putative insertion sequence IS100, identical to  
corresponding CDS from Y.pestis KIMS PCD1 (EMBL:AF053946,  
AF074612) (259 aa). fasta scores: opt: 1658 z-score:  
2693.4 E(): 0, 100.0% identity in 259 aa overlap. Similar  
to many others e.g. ISTB\_ECOLI (EMBL:X14793), istB, E.coli  
transposase for insertion sequence element IS21 (265 aa)  
(47.4% identity in 249 aa overlap). Contains PS00017  
ATP/GTP-binding site motif A (P-loop)"  
/codon\_start=1  
/transl\_table=11  
/label="YPCD1.02"  
/product="putative Amp-binding protein"  
/protein\_id="CAB54879.1"  
/db\_xref="pid:e1549592"  
/db\_xref="pid:g5832425"  
/db\_xref="gi:5832425"  
/translation="MMELQHORLMALAGOLLESLISAAPALSQOAVDQESYMFLE  
HILHEFKIARHORQAMTRMAFPAYVTFEEDTPTFGAQKQLOSIRSLSPERN  
ENIVLPGSGVKTRHATAMGTAVAGITAGTTRADLLDLUSTARQGRKTKTLOR  
GVMAPRLIIDIGILPFSQEKAKLFFQVIARYRKSAMILUTSNLPFGWDQTFAGDA  
ALTSMLDRILHSHVVOIKGSEYRLQRKRKAGVIAEANPE"  
1434. .1457  
/gene="YPCD1.02"  
/note="PS00017 ATP/GTP-binding site motif A (P-loop)"  
complement(1956. .2204)  
/partial  
/gene="YPCD1.03c"  
/note="YPCD1.03c, probable transposase remnant, len: 83  
aa; similar to several e.g. N-terminus of TRA0\_ECOLI  
(EMBL:X14793), istA, E.coli transposase for insertion  
sequence element IS21 (390 aa), similarity is interrupted  
by the adjacent IS100 element. The remainder of this CDS  
is in YPCD1.97c. This region is also similar to TR:068707  
(EMBL:AF053946, AF074612) Y.pestis KIMS putative  
transposase in PCD1 (390 aa). Contains probable  
helix-turn-helix motif at aa 19-40 (Score 1833, +5.43 SD)"  
/codon\_start=1  
/transl\_table=11  
/label="YPCD1.03c"  
/protein\_id="CAB54880.1"  
/db\_xref="pid:e1549593"  
/db\_xref="pid:g5832426"  
/db\_xref="gi:5832426"  
/translation="MUSRDEFYMIKORQOAGYIIDIAVOGCSERTVARYIUKPEPP  
ARKTRHKWKLKPFMDYIDMLRAEWNVNSEVILAIKAK"  
1956. .46271  
/note="identical to Y.pestis KIMS plasmid PCD1  
(EMBL:AF053946) from 46489 to 1955, and to Y.pestis KIMS  
plasmid PCD1 (EMBL:AF074612) from 59097 to 14563, except  
where noted"  
complement(1956. .2204)  
/gene="YPCD1.03c"  
2304. .2948  
/gene="YPCD1.04"  
2304. .2948  
/partial  
/gene="YPCD1.04"  
/note="YPCD1.04, possible transposase remnant, len: 215

aa: similar to many e.g. TR:046612 (EMBL:X78052)  
Enterobacter agglomerans IS 1222 ORB (2/6 aa). truncated  
at N-terminus"  
/codon\_start=1  
/transl\_table=11  
/label=yPCD1.04  
/protein\_id="CAB54881.1"  
/db\_xref="PID:e1549594"  
/db\_xref="PID:g5832427"  
/db\_xref="GI:5832427"  
/translation="HREGIIVHNKRYRRLYLHSLGLGVRKRRKGLATERLLPLPPA  
PILTMSMDVDMALATGRIRKICLTCDVDYTKELCYVFAFGISGVQVRIIDSLALF  
GPAIRTRPGCEETRCADPDWAFEGHVGELRILOGKPRQNEISFNGREFDELNT  
HHEFSVNRKRTISEMRQVYNCRPHSTINVTPEBFAAMKNGMSDSCGDIYTK"  
2304. 2982  
/note="similar to E.agglomerans IS 1222 (EMBL:X78052) at  
DNA level"  
complement(3014..3406)  
/gene="yPCD1.05c"  
complement(3014..3406)  
/gene="yPCD1.05c"  
note="yPCD1.05c, syce, yera, yope targeting protein, len  
130 aa; identical to corresponding CDS from Y.pestis KIM5  
PCD1 (EMBL:AF053946, AF074612) (130 aa), fasta scores:  
opt: 861 z-score: 1608.5 E()= 0, 100.0% identity in 130 aa  
overlap and to YERA\_YERPE (EMBL:M34279) from Y.pestis  
plasmid pYO19. Highly similar to TR:056910 (EMBL:Z18559)  
syce from Yersinia enterocolitica (130 aa) (99.2%  
identity in 130 aa overlap) and YERA\_YEREN (EMBL:M34278),  
yera, also from Y.enterocolitica (130 aa) (98.5% identity  
in 130 aa overlap). Similar to TR:051448 (EMBL:L27629),  
ORF1, Pseudomonas aeruginosa exoenzyme S ORF1 (116 aa)  
(44.3% identity in 115 aa overlap)"  
/codon\_start=1  
/transl\_table=11  
/label=syce  
/product="putative yope chaperone"  
/protein\_id="CAB54882.1"  
/db\_xref="PID:e1549595"  
/db\_xref="PID:g5832428"  
/db\_xref="GI:5832428"  
/translation="MYSFQATQLFQOLSLISPTPIEYGVKGFACHTTEHPVY  
QLMFLTPSLDNDDEKELISHNISQDILKFLISNDEYGVHPVLMNRQPLNSLDNNNN  
LITQLEMLVQGAERLQTSLSLSPNSFS"  
3600. 4429  
/gene="yPCD1.06"  
3600. 4429  
CDS

Query Match	62.5%	Score 979.4	DB 1	Length 70305
Best Local Similarity	99.9%	Pred. No. 3.5e-204		
Matches 980	Conservative	0	Indels	Gaps
QY 586	ATGATTAGAGCCTAGGAACAAACCACAACTTTATTTAGAGATCTACAAAAGTTAG	645		
Db 22915	ATGATTAGAGCCTAGGAACAAACCACAACTTTATTTAGAGATCTACAAAAGTTAG	22856		
QY 646	GTGGAACAACCTACTGCTGTCATGTTCTTCAGTTTAGAAGATTGGTTCAGTTAGTC	705		
Db 22855	GTGGAACAACCTACTGCTGTCATGTTCTTCAGTTTAGAAGATTGGTTCAGTTAGTC	22796		
QY 706	GATAAAATATATGATATTTTCATTTAAATATGATCCAGAAAAGATTCGAGGTTTTGGC	765		
Db 22795	GATAAAATATATGATATTTTCATTTAAATATGATCCAGAAAAGATTCGAGGTTTTGGC	22736		
QY 766	AATAGAGTATTTACTGATATATATCGAATTCGTCACAGAAAATCCTTACGTTATTTTCCACC	825		
Db 22735	AATAGAGTATTTACTGATATATATCGAATTCGTCACAGAAAATCCTTACGTTATTTTCCACC	22676		
QY 826	GAGGATACCATTTCTTAAAGCGGCTCTTTGCAACCACTGCAAAATGCGATCCAAAGCA	885		
Db 22675	GAGGATACCATTTCTTAAAGCGGCTCTTTGCAACCACTGCAAAATGCGATCCAAAGCA	22616		
QY 886	GTAAAAAGATTCCTTGAATCATCGCCGATACACATGGAATTCGGCGCTTCATGCA	945		

Db	22615	GTAAAAAGATTCCTTGAATCATCGCCGGAATACATGGAATTCGGGCGCTCATGGCA	22556
Qy	946	GTAAATCATTTCTCTTAAACCCGCGATGCATGCATGATTAATTTTGAAGTGTATGTT	1005
Db	22555	GTAAATGCATTTCTCTTAAACCCGCGATGCATGCATGATTAATTTTGAAGTGTATGTT	22496
Qy	1006	GATTCATGAATCATCATGCTGATGATCCCGTAGCAAGTTGCGTGAAGAAATTAAGCTGAGCTT	1065
Db	22495	GATTCATGAAGAAATCATCATGATGATCCCGTAGCAAGTTGCGTGAAGAAATTAAGCTGAGCTT	22436
Qy	1066	ACCGCGGAATTTAAAGATTTATTCACATTTATTCAGCCGGAATTAATTAAGCATCTGCTAGT	1125
Db	22435	ACCGCGGAATTTAAAGATTTATTCACATTTATTCAGCCGGAATTAATTAAGCATCTGCTAGT	22376
Qy	1126	AGTGCACCATTAATATTCATGATTAATTCATTAATCTCATGATTAATAATTTATATGTT	1185
Db	22375	AGTGCACCATTAATATTCATGATTAATTCATTAATCTCATGATTAATAATTTATATGTT	22316
Qy	1186	TATACAGATGAAGAGATTTTAAAGCCAGCGCAGAGTACAAAATTTCTCGAGAAAAATGCCT	1245
Db	22315	TATACAGATGAAGAGATTTTAAAGCCAGCGCAGAGTACAAAATTTCTCGAGAAAAATGCCT	22256
Qy	1246	CAAAACCACATTCAGGTGATGAGGAGCGAGAAAAAATATAGTCTCATTAAGSACTTTCTT	1305
Db	22255	CAAAACCACATTCAGGTGATGAGGAGCGAGAAAAAATATAGTCTCATTAAGSACTTTCTT	22196
Qy	1306	GGAAGTGAATAATAAAGAACCGGGCGTTGGTAACTCGAAAAACTCATACTCTTATAT	1365
Db	22195	GGAAGTGAATAATAAAGAACCGGGCGTTGGTAACTCGAAAAACTCATACTCTTATAT	22136
Qy	1366	AAAGATTAATGAATTAATCTCTCACTTTGCCACACCTGCTCGGATTAAGTCAGGCGCGCTC	1425
Db	22135	AAAGATTAATGAATTAATCTCTCACTTTGCCACACCTGCTCGGATTAAGTCAGGCGCGCTC	22076
Qy	1426	AACGACTGGTTAGCCAAAAAACCACTGAGCTGTGATTAATCAATCAAGTTTAAATCA	1485
Db	22075	AACGACTGGTTAGCCAAAAAACCACTGAGCTGTGATTAATCAATCAAGTTTAAATCA	22016
Qy	1486	GCTATTGAAGCACTGAACCGTTTCATTCAGAAATATGATTCAGTATGCAACGCTGTGCTA	1545
Db	22015	GCTATTGAAGCACTGAACCGTTTCATTCAGAAATATGATTCAGTATGCAACGCTGTGCTA	21956
Qy	1546	GATGACACGCTGTGTPAATGA 1566	
Db	21955	GATGACACGCTGTGTPAATGA 21935	

RESULT	5
LOCUS	AF053946
DEFINITION	AF053946 70504 bp DNA circular BCT 06-OCT-1998
ACCESSION	Yersinia pestis plasmid pCD1, complete plasmid sequence.
NID	AF053946
VERSION	92996222
KEYWORDS	AF053946.1 GI:2996222
SOURCE	Yersinia pestis.
ORGANISM	Yersinia pestis
REFERENCE	Bacteria: Proteobacteria: gamma subdivision; Enterobacteriaceae: Yersinia.
AUTHORS	1 (bases 1 to 70504) Hu, P., Elliott, J., McCreedy, P., Skowronski, E., Garnes, J., Kobayashi, A., Brubaker, R. and Garcia, E.
TITLE	Structural organization of virulence-associated plasmids of Yersinia pestis
JOURNAL	J. Bacteriol. 180 (19), 5192-5202 (1998)
MEDLINE	98422474
REFERENCE	2 (bases 1 to 70504)
AUTHORS	Hu, P., Elliott, J., McCreedy, P., Skowronski, E., Garnes, J., Kobayashi, A., Carrano, A.V., Brubaker, R. and Garcia, E.
TITLE	Direct Submission
JOURNAL	Submitted (16-MAR-1998) Biology and Biotechnology, Lawrence Livermore National Lab, L452, 7000 East Ave., Livermore, CA 94550

```

FEATURES             USA
source               1. 70504
                    /organism="Yersinia pestis"
                    /plasmid="pCD1"
                    /strain="K1M"
                    /db_xref="taxon:632"
                    87. 1109
CDS
    /codon_start=1
    /transl_table=11
    /product="transposase"
    /protein_id="AAC62546.1"
    /db_xref="PID:g2996223"
    /db_xref="GI:2996223"
    /translation="MTVEFVWEIKLHKQGSRAIARELGISRNTRYKRLQAKSEP
    KPYPRPAVASLDEYRDYRIORIADAHPIPAVIAIREIDOGYRGMTILRAPIR
    SLVPOGEPAVRETEBERGROMQDMGRNSPLHYFAVLGTSRLYIEFDNMR
    YDLETCIRNAFRFFGGVREVLIDNMKTIVLQADAYTGQHRHPSLMQEKGFS
    PRLCRPFAQTGRKVERMVQYTRNSFYIPLMTRLRPMGIIVDVEYANRHRGLWLDVA
    NQRHETIOARPCDRMLEEOOSMLALPPEKREYDVHLDENLVNEDKHPDLHPLSIYDS
    FCRGVA"
    1109. 1888
CDS
    /codon_start=1
    /transl_table=11
    /product="transposase"
    /protein_id="AAC62557.1"
    /db_xref="PID:g2996234"
    /db_xref="GI:2996234"
    /translation="MELQHORLMAACQQLDESISAAPALSOQAVDOENSYMDLE
    HLHEKLIARHORQAMYTMAAPAVTFEYDTEFGAPQKQSLRSISFERN
    ENIVLPGSGVGTILAIAMGEAVRAGIKVAFTHADLLQLSTAQROGRKTIQGR
    GMAFRLILDEIGYLPFSOEAKLFEQVIARXYSAMILTSNLPFGOMQDTFAGDA
    ALTSMLDRILHSHVVOIKGFSYRLQKRAKAGVIAEANPE"
    complement(1939. 2343)
CDS
    /codon_start=1
    /transl_table=11
    /product="transposase"
    /protein_id="AAC62568.1"
    /db_xref="PID:g2996245"
    /db_xref="GI:2996245"
    /translation="MTYQASERRRCRIMGISRSLIHYCPNTARDIPVVEYLQKIAHO
    YPAFGFLMFNKLKQSGLPWVKRYRYRRLIKLNFRRKGRKRLPNRHPOPLAIPLM
    NHCSVDPMDSALTDGRFRFLFNVEIVNDG"
    complement(2379. 2645)
    /gene="IcrS"
    /complement(2379. 2645)
gene
    /gene="IcrS"
    /codon_start=1
    /transl_table=11
    /product="low calcium response locus protein S homolog"
    /protein_id="AAC62579.1"
    /db_xref="PID:g2996256"
    /db_xref="GI:2996256"
    /translation="MKKARFTETQILRYLKEVEGGRHYKDVCRENGSGSEASYNMKSK
    YGGMSSDIKRNKEREENRRLKOMYASLSLDHEILKDVAAKL"
    complement(3193. 3540)
    /gene="yscM"
    /complement(3193. 3540)
CDS
    /gene="yscM"
    /codon_start=1
    /transl_table=11
    /product="Yop proteins translocation protein M homolog"
    /protein_id="AAC62586.1"
    /db_xref="PID:g2996263"
    /db_xref="GI:2996263"
    /translation="MKINTLOSLNQIITOVGHGGOAGRLTFETNPLENSHOISTAEK
    AFANVELHVKNKTAISRHDIACLPRVSNLELKQKAGEVITGTRTQLSLSDAKKL
    LEAARODTADG"
    complement(3765. 4430)
    /gene="yscL"
    /complement(3765. 4430)
CDS
    /gene="yscL"
    /codon_start=1
    /transl_table=11
    /product="Yop proteins translocation protein L homolog"
    /protein_id="AAC62596.1"
    /db_xref="PID:g2996273"
    /db_xref="GI:2996273"
    /translation="MSOTCCQGYAVQPFVQIIPNSNLSACGLRILRAEDYQSSLTTE
    ELISAAKQDAEKILADAOEYEQKOLGOMGMDARTLATLIHERTOLQOCQFRRHY
    EQQMSFVLLAVLRKILNDYDQVAMTIDVREALAVSNQKQVYVFNPKDQAGAIREQI
    AKVHKDFPEISYLEVTADARLDQGCILTEFVGILIDASIDQIALSAISTLTIGQMK
    VTE"
    complement(4376. 5005)
    /gene="yscK"
    /complement(4376. 5005)
CDS
    /gene="yscK"
    /codon_start=1
    /transl_table=11
    /product="Yop proteins translocation protein K homolog"
    /protein_id="AAC62605.1"
    /db_xref="PID:g2996282"
    /db_xref="GI:2996282"
    /translation="MMENYIRSFQLRCPAVYHLFOLPSIMRSIPLYLPQWRDRAHL
    NNAIDERSLDYEBERHGALPLQROSOLELLCRGLVLRGEAIRCVYASPLQD
    LTLVNDLETKQIIVQHELLIGPWPTHQRLPTEISRTMQLGSLAWLAHEPQO
    AMCKRLSLRLPASPSEPVLVAESORPLAOTLCHKLYKQVPTCSHLFK"
    complement(5005. 5739)
    /gene="yscJ"
    /complement(5005. 5739)
CDS
    /gene="yscJ"
    /codon_start=1
    /transl_table=11
    /product="Yop proteins translocation protein J"
    /protein_id="AAC62607.1"
    /db_xref="PID:g2996284"
    /db_xref="GI:2996284"
    /translation="MKVKTSLSTLILFLTGCKVVDYTGISQEGNEMALLRQEG
    SADKEPDKIKLILVESDVAAIDILKRQYHESSTLDQVFPKGLISPSIEEL
    ARLVAKAQEISRTLSIEDGVAVRHHVLEEDNNKGRKGAASAVFIKHADIOF
    DTYIPIQKOLVNNISIEGLAYDRISVILPVSVDVQSSHLPRNTSILSIQVSESGHL
    IGLSILILPLVNNIAQFWLQRRK"
    complement(5746. 6093)
    /gene="yscI"
    /complement(5746. 6093)
CDS
    /gene="yscI"
    /codon_start=1
    /transl_table=11
    /product="Yop proteins translocation protein I homolog"
    /protein_id="AAC62608.1"
    /db_xref="PID:g2996285"
    /db_xref="GI:2996285"
    /translation="MPNTEIQADEVITTTIEIGRAPPTDQINRPAASPDQGL
    GHSILKEYSDIQKSEFTYKSDLHKLAVASDNPDMIMQMSLIRITIQEELIATAG
    RMSQNVETLSKGG"
    complement(6094. 6591)
    /gene="yscH"
    /complement(6094. 6591)
CDS
    /gene="yscH"
    /codon_start=1
    /transl_table=11
    /product="Yop proteins translocation protein H homolog"
    /protein_id="AAC62547.1"
    /db_xref="PID:g2996224"
    /db_xref="GI:2996224"
    /translation="MTVTLNKGSITSLMSSQAVSTLOPVASELKTQLENKLSAESAE
    KTRREVLMQOYVASNPPDAVLEVALPYREALLARFGQHQSVPALDLPILRSYLOQ
    FDSFGKREAILLOVLEGIKPNESQGVPIYSELINKELMILLPSNSIVDSLHNSHQ
    IDMTQ"
    complement(6588. 6935)
    /gene="yscG"
    /complement(6588. 6935)
CDS
    /gene="yscG"
    /codon_start=1
    /transl_table=11
    /product="Yop proteins translocation protein G homolog"
```

```
/protein_id="AAC62548.1"
/db_xref="PID:82996225"
/db_xref="GI:2996225"
/translation="MKYKLNVLAEIALIGTGNHYHEANCIAEMLHLKGEBAVOLIRLSISLMNRDYSALQOQNKLAYPDLEPWALCEYRLGISLESRLNRLARSQDPRIOFVNGMRQLKT"
complement(6937..7200)
/gene="yscf"
/complement(6937..7200)
/gene="yscf"
/complement(6937..7200)
/gene="yscf"
/codon_start=1
/transl_table=11
/product="Yop proteins translocation protein F homolog"
/protein_id="AAC62549.1"
/db_xref="PID:82996226"
/db_xref="GI:2996226"
/translation="MSNFSGFRTGDIADLDAVAOQLTKPADANKAVDSIALKDKPNNPLADLOHSINKMSYIYNINSIVRSMMDLMQGIQKRP"
complement(7201..7401)
/gene="yscf"
/complement(7201..7401)
/gene="yscf"
/codon_start=1
/transl_table=11
/product="Yop proteins translocation protein E homolog"
/protein_id="AAC62550.1"
/db_xref="PID:82996227"
/db_xref="GI:2996227"
/translation="MTOLEOLHNVEIVRSITMOLEMALTKLKDMMRGDAKQYQVWQREKALESAIRIHYVAGDLK"
complement(7398..8657)
gene
Query Match 62.5%; Score 979.4; DB 2; Length 70504;
Best Local Similarity 99.9%; Pred. No. 3.5e-204;
Matches 980; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
gene
586 ATGATTAGAGCCACGACAAACCCACACATTTATTCAGATCTAGAAAAAGTTAGG 645
25311 ATGATTAGAGCCACGACAAACCCACACATTTATTCAGATCTAGAAAAAGTTAGG 25370
646 GTGGACACACTTACTGGTCATGCTTCTTCAGTTTGAAGAATGGTTCAGTACAA 705
25371 GTGGACACACTTACTGGTCATGCTTCTTCAGTTTGAAGAATGGTTCAGTACAA 25430
706 GATAAATATATAGATATTTCCATTAAATATGATCCAGAAAAGATTCGAGGTTTTGCC 765
25431 GATAAATATATAGATATTTCCATTAAATATGATCCAGAAAAGATTCGAGGTTTTGCC 25490
766 AATAGACTAATTCATGATGATTCGAATTCCTCAAGAAAATCCTAGCTTATTTCTACCC 825
25491 AATAGACTAATTCATGATGATTCGAATTCCTCAAGAAAATCCTAGCTTATTTCTACCC 25550
826 GAGGATACCATCTTAAAGCGCGCTTATGACACCAACTGCAAAATGCGATCAAGCGA 885
25551 GAGGATACCATCTTAAAGCGCGCTTATGACACCAACTGCAAAATGCGATCAAGCGA 25610
886 GTAAAAAGAGTCCCTGAATCATGCGCGAATACAAATGGGAATTCGGCGGCTCATGGCA 945
25611 GTAAAAAGAGTCCCTGAATCATGCGCGAATACAAATGGGAATTCGGCGGCTCATGGCA 25670
946 GTATATCATTTCTCTTAAACCGCCGATCGATGATGATGATTTTGAAGAGATTTGTT 1005
25671 GTATATCATTTCTCTTAAACCGCCGATCGATGATGATGATTTTGAAGAGATTTGTT 25730
1006 GATTCAATGATCATGATGGTATGCGCGTAGCAAGTTGGTGAAGATTAAGTGAAGCTT 1065
25731 GATTCAATGATCATGATGGTATGCGCGTAGCAAGTTGGTGAAGATTAAGTGAAGCTT 25790
1066 ACCGCCGAATTAAGATTTATTCAGTTATCAAGCGAATTAATTAAGCATCTGTCTAGT 1125
25791 ACCGCCGAATTAAGATTTATTCAGTTATCAAGCGAATTAATTAAGCATCTGTCTAGT 25850
1126 AATGGACCAATAATATCCATGATTAATCCATTAAATCTCATGATATAAAATTTATATGCT 1185
```

```
|||||
Db 25851 AGTGGACCAATAATATCCATGATTAATCCATTATCTCATGGATTAATAATTTATATGCT 25910
Qy 1186 TATACAGATGAAGAGATTTTAAAGCCAGCGAGATACAAATTTCTCGAGAAATGCTT 1245
Db 25911 TATACAGATGAAGAGATTTTAAAGCCAGCGAGATACAAATTTCTCGAGAAATGCTT 25970
Qy 1246 CAACACCATTCAGTGGATGGAGCGAGAAAAAATAGTCTCGATTAAGACATTTCTT 1305
Db 25971 CAACACCATTCAGTGGATGGAGCGAGAAAAAATAGTCTCGATTAAGACATTTCTT 26030
Qy 1306 GGAAGTGAATTAAGAACCGGGCGGTGGGTAATCTGAAAAACATACCTTATATAT 1365
Db 26031 GGAAGTGAATTAAGAACCGGGCGGTGGGTAATCTGAAAAACATACCTTATATAT 26090
Qy 1366 AAGATTAATTAAGATTAATCTGATTTGCGACACCTGCTCGGATTAAGTCCAGCGCGCTC 1425
Db 26091 AAGATTAATTAAGATTAATCTGATTTGCGACACCTGCTCGGATTAAGTCCAGCGCGCTC 26150
Qy 1426 AACGACTTGTTAGCCAAAAAACACTCAGCTGCTGATTAATACATCAGTTTAAATCA 1485
Db 26151 AACGACTTGTTAGCCAAAAAACACTCAGCTGCTGATTAATACATCAGTTTAAATCA 26210
Qy 1486 GCTATGAGCACTGAAACCGTTTCATTCAGAAATATGATTCGTATGCAAGCTCTGCTA 1545
Db 26211 GCTATGAGCACTGAAACCGTTTCATTCAGAAATATGATTCGTATGCAAGCTCTGCTA 26270
Qy 1546 GATGACACGCTGCTGTAATGA 1566
Db 26271 GATGACACGCTGCTGTAATGA 26291
|||||
RESULT 6
AF074612 AF074612 70559 bp DNA circular BCT 03-NOV-1998
LOCUS Yersinia pestis Plasmid pCD1, complete plasmid sequence.
DEFINITION AF074612
ACCESSION AF074612
VERSION 93822037
NID AF074612.1 GI:3822037
KEYWORDS Yersinia pestis.
SOURCE Yersinia pestis.
ORGANISM Yersinia pestis.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Yersinia.
REFERENCE 1 (bases 1 to 70559)
AUTHORS Perry,R.D., Straley,S.C., Fetherston,J.D., Rose,D.J., Gregor,J. and Blattner,F.R.
TITLE DNA sequencing and analysis of the low-Ca2+-response plasmid pCD1 of Yersinia pestis KIM5
JOURNAL Infect. Immun. 66 (10), 4611-4623 (1998)
MEDLINE 98427122
REFERENCE 2 (bases 1 to 70559)
AUTHORS Perry,R.D., Straley,S.C., Fetherston,J.D., Rose,D.J., Gregor,J. and Blattner,F.R.
TITLE Direct Submission
JOURNAL Submitted (25-JUN-1998) Microbiology and Immunology, University of Kentucky, MS415 Medical Center, Lexington, KY 40536-0084, USA
COMMENT Unique identifiers, starting at Y0001, are assigned to each gene of pCD1.
FEATURES
source Location/Qualifiers
1..70559
/organism="Yersinia pestis"
/plasmid="pCD1"
/strain="KIM5"
/db_xref="taxon:632"
57..368
/gene="Y0001"
57..368
/gene="Y0001"
/gene="Y0001"
/note="103: 43 pct identical (0 gaps) to 100 residues of an approx. 200 aa protein GENPEPT: g11537126, orf_c198 Escherichia coli"
/codon_start=1
```



```

gene
    /transl_table=11
    /product="unknown"
    /protein_id="AAC69758.1"
    /db_xref="PID:g3822038"
    /db_xref="GI:3822038"
    /translation="MHQSRGAAGSRTSLIMROSGYNNVWRILARLRECGIAROP
    CKPRRGEVNSLASPDILKROFKPSERPWRVSGTISITKNGCMYIALYIDLSFHW
    665. .1033
    /gene="nuc"
    /note="Y0002"
    665. .1033
    /gene="nuc"
    /codon_start=1
    /transl_table=11
    /product="endonuclease"
    /protein_id="AAC69759.1"
    /db_xref="PID:g3822039"
    /db_xref="GI:3822039"
    /translation="MDTKLLQHTPIGTWYDYPVPTKSGCKRLRCPDFVHYRMDL
    VNAQIPVPTVNSFKALHDKVIIVDGKNTQMGSEFNQAQVNSENVLLIMWDFVYVQ
    AYLQWOSRWNGTDMRSSY"
    1171. .1425
    /gene="repB"
    /note="Y0003"
    1171. .1425
    /gene="repB"
    /codon_start=1
    /transl_table=11
    /product="replication protein"
    /protein_id="AAC69760.1"
    /db_xref="PID:g3822040"
    /db_xref="GI:3822040"
    /translation="MSQIENAVTSSSKRAYRKGNPLTGAQKORMSVSRKKEHKAIV
    FIONLKNELLQICDSGLTOTEMIERWIQREKARTNA"
    complement(1560. .1649)
    /note="antisense RNA"
    /product="copA"
    1667. .1741
    /gene="tap"
    /note="Y0004"
    1667. .1741
    /gene="tap"
    /codon_start=1
    /transl_table=11
    /product="repa translation protein"
    /protein_id="AAC69761.1"
    /db_xref="PID:g3822041"
    /db_xref="GI:3822041"
    /translation="MFRKVYLLRLLLPLCNISAGRCD"
    1734. .2600
    /gene="repa"
    /note="Y0005"
    1734. .2600
    /gene="repa"
    /codon_start=1
    /transl_table=11
    /product="replication protein"
    /protein_id="AAC69762.1"
    /db_xref="PID:g3822042"
    /db_xref="GI:3822042"
    /translation="MTNHOALFTHHYROYKNNPPEFTPREGKTLPEFCRKLMAKAEG
    TSPRDSMVAFAARSLSLRHMPPLLRRAIDALOGCFHFDPLANRORSITLAI
    ECGLATSESGSLSTRATRALPFLSELCTITYOTIEYDIOICNTPTDITTPALFSA
    LVSDVAVAARSRVEMNQOREKORLPRLTMDLAIAMRPFVERERSYOTEKRAH
    GKRAARADVLRTRDIEAIYVNRDLTREIAGRFVGNLDAVRREKARVKERMIMSR
    NNNYIRLATGAT"
    1741. .2935
    /note="ori"
    complement(3427. .3645)
    /gene="Y0006"
    complement(3427. .3645)
    /gene="Y0006"

gene
    /note="Orf5 (f72): 42 pct identical (0 gaps) to 33
    residues of an approx. 216 aa protein GENPEPT: g1|2055297,
    Imp2 Xenopus laevis"
    /codon_start=1
    /transl_table=11
    /product="unknown"
    /protein_id="AAC69763.1"
    /db_xref="PID:g3822043"
    /db_xref="GI:3822043"
    /translation="MRSPLAGCSVYVNVVYVTLHNDIHKSAYKHKRLYHVVLGQASS
    AQPFSMHPRGKMAPADINNSVDNMR"
    4758. .5186
    /gene="Y0008"
    4758. .5186
    /gene="Y0008"
    /note="Orf1 (o142): 31 pct identical (1 gap) to 48
    residues of an approx. 104 aa protein GENPEPT: g1|2149940,
    orf1 P. syringae"
    /codon_start=1
    /transl_table=11
    /product="unknown"
    /protein_id="AAC69764.1"
    /db_xref="PID:g3822044"
    /db_xref="GI:3822044"
    /translation="MINTFTPELLPKIASHFGDLKSODEYGLCELLANDRVYIMLRA
    DEILNRLTLLEPILIGFSGPARSAASOLFQCSYNALNKGDPGFAMSEELIAFVKHL
    SLDELNVENSKEIANYDWMLSLVSLPAETAPSYISYISG"
    5204. .7402
    /gene="ypka"
    /note="Y0009"
    5204. .7402
    /gene="ypka"
    /codon_start=1
    /transl_table=11
    /product="secreted protein kinase"
    /protein_id="AAC69765.1"
    /db_xref="PID:g3822045"
    /db_xref="GI:3822045"
    /translation="MKSVKIMGTMPSSISLAKAHERISOHMONPVGELNIGKKRYRII
    DNOYLRLNPHSGESLPREGYCKITFGKMFNPSTARNITDILHAOKTTSOELSRDIPN
    ALSNLFGKRPOTLEPLGKGBEPLSGADPLGEMRAETDKFREGESHISILETKDKOHL
    VAKIERSIAEGHLEAELEAVKHIYKTGKHPNLANVGMVAVPYGKNKEALLMDEVD
    GWRCSDTLRLTLADMSKOGKINSEAYWGTIFARILDLVNHAKAGVYHNDIKGVN
    VEDRASGEVVIDGLHSRSGEOPKFTESKRAELGVNLSAKSDVFLVYSLIHL
    CIEGEEKNEPKIPNOGLDRTTSEPAHMDNGVYIHRGPIAGVETATYRPTIDIGS
    ADSRPSNEARLHLELDGTIDESAKOILLKDTITGMSPLSDVYRITTPKKRLBELSD
    LLRHLSSAATKQLDMGCVLSDUTMLVALDKAREGVDKQDKSPNSLLKTYRVI
    EDYVKGREGDKNSSTEVSPIYHRNFMLSIYEPSELQIOKHLDTSHFSIGSLVRAH
    KHLLETLEVLTLSOOGPVSSEYTGFLNRLTEAKITLSOOLNTLOOQSAKAOISII
    LINSRGMAVDAQSILORPDSTOPVAVFGTEQVYAIHROMAAMAAATLLOEVSFTDD
    MRNFTVDSIPILLIDGRSSLMDDELHVQREKRLRLTIIAERLNRLERNM"
    7798. .8664
    /gene="yopJ"
    /note="Y0010"
    7798. .8664
    /gene="yopJ"
    /codon_start=1
    /transl_table=11
    /product="targeted effector protein"
    /protein_id="AAC69766.1"
    /db_xref="PID:g3822046"
    /db_xref="GI:3822046"
    /translation="MIGPISINISGSLSEKETSLSLNEKLKIIITOLEIDISDGSM
    FHKYSRNDVEVMPALVIOANKIPENKLNLYVSPDLSTETKVINENGVRSSFTIIN
    MGEQGHISVTDYKHNGKSLIFEPANFSMGPAWLARTKVAIERYOLDPCHFQFEM
    VEMDIORSSSECGIFSLAKKLTIEBDSLIKIHEDIKILSGEENLPHDKIDPLV
    PVTFFYKTHOGKKRLNXYLNTNPQGVGVYVKNKNETIVNFRDNKNSIYDGKELSVYHK
    KRIAEYTKLV"
    9858. .10064
    /gene="Y0012"
    9858. .10064
    /gene="Y0012"
    /note="Orf8: 45 pct identical (0 gaps) to 68 residues of an

```

	approx. 560 aa protein GENEPT: g1:2182435, hypothetical"
	/codon_start=1
	/transl_table=1
	/product="unknown"
	/protein_id="AAC69767.1"
	/db_xref="PID:93822047"
	/db_xref="GI:3822047"
	/translation="MAYGRNMLFAGSLRAGQRMASILSLFTAKLNGHPYWLRLV
gene	LRLPLTPNSQLNALPLPAENRS"
	10347. .11753
	/gene="yopH"
	/note="Y0013"
CDS	10347. .11753
	/gene="yopH"
	/codon_start=1
Query Match	62.5%: Score 979.4; DB 2; Length 70559;
Best Local Similarity	99.9%: Pred. No. 3.5e-204;
Matches 980; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
QY 586	ATGATTAGAGCTTACGACAAACACACATTTTATGAGATCTAGAAAAAGTTAGG 645
DB 37919	ATGATTAGAGCTTACGACAAACACACATTTTATGAGATCTAGAAAAAGTTAGG 37918
QY 646	GTGGAACACTTACTGCTCATGCTTCTCAGTTTGAAGAATTGGTTCAGTTAGTCAA 705
DB 37979	GTGGAACACTTACTGCTCATGCTTCTCAGTTTGAAGAATTGGTTCAGTTAGTCAA 38038
QY 706	GATAAAAATATAGATATTTCCATTAAATATGATCCCAAAAAGATTGCGAGTTTTC 765
DB 38039	GATAAAAATATAGATATTTCCATTAAATATGATCCCAAAAAGATTGCGAGTTTTC 38098
QY 766	AATAGAGTAATTAAGATATGATATGCAATTTGCTCAAGAAATCTAGCTTATTTTC 825
DB 38099	AATAGAGTAATTAAGATATGATATGCAATTTGCTCAAGAAATCTAGCTTATTTTC 38158
QY 826	GAGATACCATTCTTAAAGGGGCTCATATGACAAACCACTGCAAAATGCGATCAAGCA 885
DB 38159	GAGATACCATTCTTAAAGGGGCTCATATGACAAACCACTGCAAAATGCGATCAAGCA 38218
QY 886	GTAAGAAGTCTTGAATCATCGCCGAATACAAATGGAATTGCGGCGTTCATGGCA 945
DB 38219	GTAAGAAGTCTTGAATCATCGCCGAATACAAATGGAATTGCGGCGTTCATGGCA 38278
QY 946	GTAATGATTTCTTTTAAACCGCGATGATGATGATATTTTGAAGTATGTT 1005
DB 38279	GTAATGATTTCTTTTAAACCGCGATGATGATGATATTTTGAAGTATGTT 38338
QY 1006	GATTCATGATCATCATGCTGATGCGGTAGCAAGTTGCGTGAAGATTTAGCTGAGCTT 1065
DB 38339	GATTCATGATCATCATGCTGATGCGGTAGCAAGTTGCGTGAAGATTTAGCTGAGCTT 38398
QY 1066	ACCGCCGAATTAAGATTATTCAGTTATTCAGCCGAATTAATACCACTGCTAGT 1125
DB 38399	ACCGCCGAATTAAGATTATTCAGTTATTCAGCCGAATTAATACCACTGCTAGT 38458
QY 1126	AGTGGCACCATTAATTCATGATTAATCATTAATCATGATTAATTAATTAATGCT 1185
DB 38459	AGTGGCACCATTAATTCATGATTAATCATTAATCATGATTAATTAATTAATGCT 38518
QY 1186	TATACAGATGAAGATTTTAAAGCCAGCGAGTCAAAATTTCTGAGAAAAATGCGCT 1245
DB 38519	TATACAGATGAAGATTTTAAAGCCAGCGAGTCAAAATTTCTGAGAAAAATGCGCT 38578
QY 1246	CAAACACCATTCAGTGTGATGGAGGAGGAAAAAATAGCTCGATTAAGACTTTCTT 1305
DB 38579	CAAACACCATTCAGTGTGATGGAGGAGGAAAAAATAGCTCGATTAAGACTTTCTT 38638
QY 1306	GGAAGTGAATTAAGAACCGGCGCTTGGGTATCTGAAAAACTCATCTTATAT 1365
DB 38639	GGAAGTGAATTAAGAACCGGCGCTTGGGTATCTGAAAAACTCATCTTATAT 38698
QY 1366	AAAGATTAATTAATGATTAATCTACTTTCACACCACTGCTCGATTAAGTCCAGGCGCTC 1425

DB 38699	AAAGTATTAATGAATTAATCTACCTTGGCCACACCTGCTCGGATTAAGTCCAGGCCCTC 38758
QY 1426	AACGACTTGGTTAGCCAAAAAACACACACAGCTGCTGATATATACATGACGTTTAATCA 1485
DB 38759	AACGACTTGGTTAGCCAAAAAACACACAGCTGCTGATATATACATGACGTTTAATCA 38818
QY 1486	GCTATTAAGACACGTAACCGCTTTCATTCAGAAATATGATGATTCAGTGAACGCTGCTA 1545
DB 38819	GCTATTAAGACACGTAACCGCTTTCATTCAGAAATATGATGATTCAGTGAACGCTGCTA 38878
QY 1546	GATGACAGCTGCTGTAATGA 1566
DB 38879	GATGACAGCTGCTGTAATGA 38899
RESULT 7	
A46411	A46411 1014 bp DNA PAT 07-MAR-1997
LOCUS	Sequence 1 from Patent W09524475.
DEFINITION	A46411
ACCESSION	A46411
NID	92300612
VERSION	A46411.1 GI:2300612
KEYWORDS	
SOURCE	Yersinia pestis.
ORGANISM	Yersinia pestis
REFERENCE	1. (bases 1 to 1014)
AUTHORS	Titchell, R.W., Williamson, E.D. and Leary, S.E.
TITLE	VACCINE COMPOSITIONS
JOURNAL	Patent: WO 9524475-A 1 14-SEP-1995;
COMMENT	SECR DEFENCE BRIT (GB)
FEATURES	Other publication AU 1853995 950925.
source	Location/Qualifiers
	1. 1014
	/organism="Yersinia pestis"
	/db_xref="taxon:632"
CDS	1. 990
	/note="unnamed protein product"
	/codon_start=1
	/transl_table=11
	/protein_id="CA402893.1"
	/db_xref="PID:e306155"
	/db_xref="PID:92300613"
	/db_xref="GI:2300613"
	/translation="MSEFIRAYRQNPQHPFIEDLEKYRVEQLTGHSSVLEIYQVWD
	KNTDISIKYPRDSEYFANRYITTDIELLKLTLATFLPEDALIKGHTDNOLOKIK
	RKAEFLSSPNTQWELRAFYAVHFSLTADRIDDDILKYVDSMHHGDARSKLREEL
	LELTAEIKIYSVIAQAIKHLSSSGTINIDKSIINMDKNLYGTDEIFKASAEYKI
	LEKMPQTTIVDGESEKIVSIKDFLSENRGTALNKLKSYSYNRDNDELSHFATTC
	SDKSRPLNDIVSQKTTOLSDTTSRFSATIEALNRFIQKIDSVQRLDDPTSGK"
BASE COUNT	346 a 181 c 201 g 286 t
ORIGIN	
Query Match	62.4%: Score 976.8; DB 5; Length 1014;
Best Local Similarity	99.8%: Pred. No. 1.4e-203;
Matches 978; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY 587	TGATTAGAGCTTACGACAAACACACATTTTATGAGATCTAGAAAAAGTTAGG 646
DB 11	TCATTAGAGCTTACGACAAACACACATTTTATGAGATCTAGAAAAAGTTAGG 70
QY 647	TGGAACCACTTACTGCTCATGCTTCTCAGTTTGAAGAATTGGTTCAGTTAGTCAA 706
DB 71	TGGAACCACTTACTGCTCATGCTTCTCAGTTTGAAGAATTGGTTCAGTTAGTCAA 130
QY 707	ATAAAAAATATGATATTTCCATTAAATATGATCCCAAAAAGATTGCGAGGTTTTC 766
DB 131	ATAAAAAATATGATATTTCCATTAAATATGATCCCAAAAAGATTGCGAGGTTTTC 190
QY 767	ATAGATTAATTAATGATTAATCTACTTTCACACCACTGCTCGATTAAGTCCAGGCGCTC 826



QY 1247 AACACACCATTCAGTGTGATGGAGCGAAGAAAAATAGTCTCGATAAAGACTTCTTG 1306  
|||||  
Db 671 AACACACCATTCAGTGTGATGGAGCGAAGAAAAATAGTCTCGATAAAGACTTCTTG 730  
QY 1307 GAAGTGAATTAAGAAAGACCGGGCGTGGTATCTGAAAACTACTACTATATA 1366  
|||||  
Db 731 GAAGTGAATTAAGAAAGACCGGGCGTGGTATCTGAAAACTACTACTATATA 790  
QY 1367 AAGATAATTAATGAATTAATCTACTTTGCCACACCTCTCGGATTAAGTCCAGGCCGTCA 1426  
|||||  
Db 791 AAGATAATTAATGAATTAATCTACTTTGCCACACCTCTCGGATTAAGTCCAGGCCGTCA 850  
QY 1427 AACACTGGTATGCCAAAAACAACATCAGCTGTCTGATTAATACATCACTTTAATTCAG 1486  
|||||  
Db 851 ACGACTGGTATGCCAAAAACAACATCAGCTGTCTGATTAATACATCACTTTAATTCAG 910  
QY 1487 CTATGTAGACACTGAACCGTTTCATTCAGAAATATGATTCAGTATCAACAGTCTGCTAG 1546  
|||||  
Db 911 CTATGTAGACACTGAACCGTTTCATTCAGAAATATGATTCAGTATCAACAGTCTGCTAG 970  
QY 1547 ATGACAGCTCTGTAAATGA 1566  
|||||  
Db 971 ATGACAGCTCTGTAAATGA 990

RESULT 9  
A46413 1014 bp DNA PAT 07-MAR-1997  
LOCUS Sequence 3 from Patent WO9524475.  
DEFINITION A46413  
ACCESSION 92300614  
NID A46413.1 GI:2300614  
VERSION  
KEYWORDS  
SOURCE Yersinia pestis.  
ORGANISM Yersinia pestis  
Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;  
Yersinia.  
REFERENCE 1 (bases 1 to 1014)  
AUTHORS Tibball,R.W., Williamson,E.D. and Leary,S.E.  
TITLE VACCINE COMPOSITIONS  
JOURNAL Patent: WO 9524475-A 3 14-SEP-1995;  
COMMENT SECR DEFENCE BRIT (GB)  
FEATURES  
source  
1..1014  
/organism="Yersinia pestis"  
/db\_xref="taxon:632"  
1..987  
/note="unnamed protein product; Protein sequence is in  
conflict with the conceptual translation"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA02894.1"  
/db\_xref="PID:e306384"  
/db\_xref="PID:g2300615"  
/db\_xref="GI:2300615"  
/translation="GIPGRAYEQNPQHFIEDLEKYRVDELTHGSSVYLEVQVYKD  
KNIDIKYDPRKDESEFANRVITDDIELKILAFLEDALIKGHDYNOQNGIK  
RVEFLESPNTQMELEAFMAVHFSILTADRIDDLIKYVDSNMHGDARSLREEL  
AELTAEKIVSVIOAEINKHLSSTGINTIHDSINIMDKVLVGTDELEFKASAEYKI  
LEKMPOTTIQVDSSEKKIYSIKPILSEKRRKTRALNKRYSYKNDNNEISFATTC  
SDKSRPLNDLVSKITQLOSLDTSIRFNSAIEALNRFQKIDSVQMRLLDPTSGK"

BASE COUNT 343 a 184 c 205 g 282 t  
ORIGIN

Query Match 62.2%; Score 974.8; DB 5; Length 1014;  
Best Local Similarity 99.8%; Pred. No. 3.8e-203;  
Matches 976; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 589 ATTGAGCCTACGAAACCAACCAACATTTTATGAGATAGAAAAAGTTAGGCTG 648  
|||||  
Db 13 ATTGAGCCTACGAAACCAACCAACATTTTATGAGATAGAAAAAGTTAGGCTG 72

QY 649 GAACAACTTACTGCTCATGTGTTCTTCAAGTTTAAAGAAATGTTAGTCAAGAT 708  
|||||  
Db 73 GAACAACTTACTGCTCATGTGTTCTTCAAGTTTAAAGAAATGTTAGTCAAGAT 132  
QY 709 AAAAATTAATGATTTTCCATTAATATGATCCGAGAAAGATCGGAGGTTTGGCAAT 768  
|||||  
Db 133 AAAAATTAATGATTTTCCATTAATATGATCCGAGAAAGATCGGAGGTTTGGCAAT 192  
QY 769 AGAGTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 828  
|||||  
Db 193 AGAGTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 252  
QY 829 GATACCACTTCTTAAAGCGGTCAATATGACACCAACTGCAAAATGCGATCAAGCAGTA 888  
|||||  
Db 253 GATGCCATCTTAAAGCGGTCAATATGACACCAACTGCAAAATGCGATCAAGCAGTA 312  
QY 889 AAGAGTTCCTTAATCATGCGGAGATGATGATGATGATGATGATGATGATGATGATGAT 948  
|||||  
Db 313 AAGAGTTCCTTAATCATGCGGAGATGATGATGATGATGATGATGATGATGATGATGAT 372  
QY 949 ATGCATTTCTTTTAAACCGCGGATCGATGATGATGATGATGATGATGATGATGATGAT 1008  
|||||  
Db 373 ATGCATTTCTTTTAAACCGCGGATCGATGATGATGATGATGATGATGATGATGATGAT 432  
QY 1009 TCAATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1068  
|||||  
Db 433 TCAATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 492  
QY 1069 GCGGAATTAAGATTTATTCAGTTATTCAGGCGGAATTAAGATGATGATGATGATGAT 1128  
|||||  
Db 493 GCGGAATTAAGATTTATTCAGTTATTCAGGCGGAATTAAGATGATGATGATGATGAT 552  
QY 1129 GGCACCAATAATATCCATGATGATGATGATGATGATGATGATGATGATGATGAT 1188  
|||||  
Db 553 GGCACCAATAATATCCATGATGATGATGATGATGATGATGATGATGATGATGAT 612  
QY 1189 ACAAGTGAAGAGATTTTAAAGCGGACGAGTACAAAATTCGAGAAATGCTTCAA 1248  
|||||  
Db 613 ACAAGTGAAGAGATTTTAAAGCGGACGAGTACAAAATTCGAGAAATGCTTCAA 672  
QY 1249 ACCACCTTCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1308  
|||||  
Db 673 ACCACCTTCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 732  
QY 1309 AGTGAATTAAGAAAGACCGGGCGTGGTATCTGAAAACTCATACCTTTAATATAA 1368  
|||||  
Db 733 AGTGAATTAAGAAAGACCGGGCGTGGTATCTGAAAACTCATACCTTTAATATAA 792  
QY 1369 GATTAATTAATGAATTAATCTCACTTGGCCACCACCTGCTGGATAAGTCCAGGCGTCAAC 1428  
|||||  
Db 793 GATTAATTAATGAATTAATCTCACTTGGCCACCACCTGCTGGATAAGTCCAGGCGTCAAC 852  
QY 1429 GACTGGTTGCCAAAAAAACAACTCAGCTGTCTGATTAATTCATCAGCTTTAATACGT 1488  
|||||  
Db 853 GACTGGTTGCCAAAAAAACAACTCAGCTGTCTGATTAATTCATCAGCTTTAATACGT 912  
QY 1489 ATTGAACACATGAACCGTTTCATTCAGAAATATGATTCAGTATGCAACGTCTGATAG 1548  
|||||  
Db 913 ATTGAACACATGAACCGTTTCATTCAGAAATATGATTCAGTATGCAACGTCTGATAG 972  
QY 1549 GACACGCTCTGTAAATGA 1566  
|||||  
Db 973 GACACGCTCTGTAAATGA 990

RESULT 10  
A56795 1014 bp DNA PAT 03-MAR-1998  
LOCUS Sequence 3 from Patent WO9628551.  
DEFINITION A56795  
ACCESSION 93712810  
NID A56795.1 GI:3712810  
VERSION

```

KEYWORDS      Yersinia pestis.
              Yersinia pestis.
ORGANISM      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
              Yersinia.
REFERENCE     1 (bases 1 to 1014)
AUTHORS       Ribball,R.W., Williamson,E.D., Leary, Sophie,E.C., Oyston,P.C.,
              Bennett and Alice,M.
TITLE         VACCINES FOR PLAGUE
              Patent: WO 9628551-A 3 19-SEP-1996;
              SECR DEFENCE (GB)
COMMENT       Other publication ZA 9602036 960716
              Other publication AU 4951196 961002.
FEATURES
  source      1..1014
              location/Qualifiers
              /organism="Yersinia pestis"
              /db_xref="taxon:632"
              <1..990
              /note="unnamed protein product"
              /codon_start=1
              /transl_table=1
              /protein_id="CA03420.1"
              /db_xref="PID:e1255218"
              /db_xref="PID:g3712811"
              /db_xref="GI:3712811"
              /translation="GIGIRAYEQNPQHFIEDLEKRVQEQLTGSSVLEELIYLVKD
KNIDISIKYDPRKDSVEANRANVITDIEELKRIIAYFPEDAIILKGHYNOLONGIK
RVEFLESPTOMELRAFMAVMEHFSITADRIDDDILKVIYDSMHHGDARSKIRELL
AELIAELIKYVIOIAEIKHLSGSGTINIRKSTINLMDKNLYGTYDEITKASREYTI
LEKMPQTTIOVDSSEKIVSIKDFLGSSENKRTGALGNKNSITKNKDNNEIHSATTC
SDKRPPLNDIVSQTTTOLDSTISFRNSIEALNRFIOKRYDSVMORLLDPTSG"
BASE COUNT   343 a      184 c      205 g      282 t
ORIGIN

```

Query Match	62.2%	Score 974.8:	DB 5:	Length 1014:
Best Local Similarity	99.8%:	Pred. No. 3.8e-203:		
Matches 976:	Conservative	0:	Matches 2:	Indels 0:
				Gaps 0:
QY 589	ATTAGAGCCTAGAACCAAAACCACAACTTTATTATGAGATCTAGAAAAGTTGGGTG	648		
Db 13	ATTGAGGCTAGAACCAAAACCACAACTTTATTATGAGATCTAGAAAAGTTGGGTG	72		
QY 649	GAACAACCTACTGCGATCGTCTTCAGTTTATAGAAGATTGGTCAGTTAGTCAAGAT	708		
Db 73	GAACAACCTACTGCGATCGTCTTCAGTTTATAGAAGATTGGTCAGTTAGTCAAGAT	132		
QY 709	AAAAATATGATATTTCATTTCAATTAATATGATCCACAGAAAAGATTCGAGGTTTTCCAAT	768		
Db 133	AAAAATATGATATTTCATTTCAATTAATATGATCCACAGAAAAGATTCGAGGTTTTCCAAT	192		
QY 769	AGAGTAAATTACTGATGATATCGAATTGCTCAAGAAAATTCAGCTATTTTCTACCCGAG	828		
Db 193	AGAGTAAATTACTGATGATATCGAATTGCTCAAGAAAATTCAGCTATTTTCTACCCGAG	252		
QY 829	GATFACCATTCTTAAAGCGGTCATTATGACAACAACCAATGCAAAATGGCATCAGCGAGTA	888		
Db 253	GATFACCATTCTTAAAGCGGTCATTATGACAACAACCAATGCAAAATGGCATCAGCGAGTA	312		
QY 889	AAAGAGTTCCTTGATCATCGCCGGAATACACAATGGGAATTCGGGGCGTTCATGCACTA	948		
Db 313	AAAGAGTTCCTTGATCATCGCCGGAATACACAATGGGAATTCGGGGCGTTCATGCACTA	372		
QY 949	ATGCAATTCCTCTTTAACCGCGGATGATGCAATGATGATTTTGAAGTGATTTCTGAT	1008		
Db 373	ATGCAATTCCTCTTTAACCGCGGATGATGCAATGATGATTTTGAAGTGATTTCTGAT	432		
QY 1009	TCAATGAATCATCATGCTGATGCCCCGTAGACAAGTTGCGTGAAGAATTTAGCTAGCTTACC	1068		
Db 433	TCAATGAATCATCATGCTGATGCCCCGTAGACAAGTTGCGTGAAGAATTTAGCTAGCTTACC	492		
QY 1069	GCCGAATTAAGAGTTTATTCAGTTATTCAGCCGAATTAATTAAGCATCTGTCTAGTAGT	1128		

Db	493	GGCGAATTAAAGATTATTCACGTATTATTCAGCCGCAAAATTAAAGCATCTGCTACGATG	552
Qy	1129	GGCACCATAAATATTCATGATATAATTCATTAACTCATGATGATAAAATTTATATAGTTAT	1188
Db	553	GGCCACCAATATATTCATGATATAATTCATTAACTCATGATGATAAAATTTATATAGTTAT	612
Qy	1189	ACGATGAAGATTTTAAAGCAGCCGAGTCAAAATCTTCGAGAAATAGCCCTCA	1248
Db	613	ACGATGAAGATTTTAAAGCAGCCGAGTCAAAATCTTCGAGAAATAGCCCTCA	672
Qy	1249	ACCACCATTCAGATGGATGGAGCCGAGAAAAAATAGTCTGATAAAGGACTTCTTGA	1308
Db	673	ACCACCATTCAGATGGATGGAGCCGAGAAAAAATAGTCTGATAAAGGACTTCTTGA	732
Qy	1309	AGTAGAATTAAGAAACCGGGGCGTGGGTATCTGAAAACTATACTCTTATATATA	1368
Db	733	AGTAGAATTAAGAAACCGGGGCGTGGGTATCTGAAAACTATACTCTTATATATA	792
Qy	1369	GATATATATGAATTTATCTCACTTGGCACACCTCTCGATTAAGTCAAGCCGCTAAC	1428
Db	793	GATATATATGAATTTATCTCACTTGGCACACCTCTCGATTAAGTCAAGCCGCTAAC	852
Qy	1429	GACTTGGTTAGCCAAAAACAACACTAGCTGTCTGATTTACATCAGCTTTTAAATACGT	1488
Db	853	GACTTGGTTAGCCAAAAACAACACTAGCTGTCTGATTTACATCAGCTTTTAAATACGT	912
Qy	1489	ATTGAACACTGAACCGTTCAATCGAAATATGATAGTATGCAACGCTGCTGAT	1548
Db	913	ATTGAACACTGAACCGTTCAATCGAAATATGATAGTATGCAACGCTGCTGAT	972
Qy	1549	GACACGCTCTGTAATGA	1566
Db	973	GACACGCTCTGTAATGA	990

RESULT	11				
IPETIVANT					
LOCUS	981 bp	DNA		BCT	13-MAR-1997
DEFINITION	Y.pseudotuberculosis V antigen gene.				
ACCESSION	X96802				
NID	91405834				
VERSION	X96802.1	GI:1405834			
KEYWORDS	V antigen.				
SOURCE	Yersinia pseudotuberculosis.				
ORGANISM	Yersinia pseudotuberculosis.				
REFERENCE	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae				
AUTHORS	Yersinia.				
TITLE	1 (bases 1 to 981)				
JOURNAL	Roggenkamp, A., Gelger, A.M., Leitnitz, L., Kessler, A. and				
MEDLINE	Heesemann, U.				
REFERENCE	Passive immunity to infection with Yersinia spp. mediated by				
AUTHORS	anti-recombinant V antigen is dependent on polymorphism of V				
TITLE	antigen				
JOURNAL	Infect. Immun. 65 (2), 446-451 (1997)				
MEDLINE	97162308				
REFERENCE	2 (bases 1 to 981)				
AUTHORS	Roggenkamp, A.				
TITLE	Direct Submission				
JOURNAL	Submitted (22-MAR-1996) A. Roggenkamp, Hygiene and Microbiology				
	University of Wuerzburg, Josef-Schneider Strasse 2, 97080				
	Wuerzburg, FRG				
FEATURES	Location/Qualifiers				
source	1..981				
	/organism="Yersinia pseudotuberculosis"				
	/sub_species="type I"				
	/db_xref="taxon:653"				
	1..981				
	/codon_start=1				
	/transl_table=11				
	/product="V antigen"				
	/protein_id="CAA6594.1"				
	/db_xref="PID:e239989"				
	/db_xref="PID:g1405835"				
CDS					

```

/db.xref="GI:1405835"
/db.xref="SWISS-PROT:P3394"
/translation="MIRAYEONPQHFTEDLEKVEOQLTGSSVLEELVOLYVDKNNI
DISIKYDPKRDSEVFANRVITDDLELKLILAYLPEDAILKGHYNOQLONGIKRKV
EFLESPNTOWELNAPFVHIFSLTAPRIDDDILKIVDYSNMHHGDAKSLREJLAE
TAEIKYISIOAEINKHLSGGGTINIDKSNIMDKNINQYTDIEIKASAEYKILKE
MPOTIOEGEETKEIVSKNLFLESEKRTGKGLKQDSYVKNKNNELSHATTCDSK
SRPLNDLVSKOTLTQSLDITSFNSAILEANLFIOKIDYSVMORLLDITSGK"
BASE COUNT      341 a      174 c      191 g      275 t
ORIGIN

```

Query Match	61.1%	Score 957	DB 1	length 981
Best Local Similarity	98.5%	Pred. No.	2.9e-199	
Matches 966	Conservative	0	Mismatches 15	Indels 0
			Gaps	0

QY	386	ATGATTAAAGCCCTACGACCAAAACCACAACTTTTATATGAGACTCGAAAAAGTTAGG	645
Db	1	ATGATTAAAGCCCTACGACCAAAACCACAACTTTTATGAGACTCGAAAAAGTTAGG	60
QY	646	GTGGAACAACCTTACGTGTCATGCTCTTCAGTTTATAGAAGAAATGTGTCAGTTACTCAA	705
Db	61	GTGGAACAACCTTACGTGTCATGCTCTTCAGTTTATAGAAGAAATGTGTCAGTTACTCAA	120
QY	706	GATAAAATATAGATATTTCATTTCAATTAATATGATCCAGAAAAAGATTCGAGGTTTGGC	765
Db	121	GATAAAATATAGATATTTCATTTCAATTAATATGATCCAGAAAAAGATTCGAGGTTTGGC	180
QY	766	AATAGAGTAATTTACTGATGATATGCAATTTGCTCAAGAAAATCTTAGCTTATTTTCTACCC	825
Db	181	AATAGAGTAATTTACTGATGATATGCAATTTGCTCAAGAAAATCTTAGCTTATTTCTACCC	240
QY	826	GAGGATACCATTCTTAAAGGGCGGTATATGCAACCAATGCAAAATGGCATCAAGGA	885
Db	241	GAGGATGCCATTCTTAAAGGGCGGTATATGCAACCAATGCAAAATGGCATCAAGGA	300
QY	886	GTAAGAAGATTCCTTGAATCATCGCCGGAATACACATGGGAATTCGGGCGTTCAAGCA	945
Db	301	GTAAGAAGATTCCTTGAATCATCGCCGGAATACACATGGGAATTCGGGCGTTCAAGCA	360
QY	946	GTAATGCATTTCTCTTTAAACCGCGCATGATATGATGATATTTTGAAGTATTTGT	1005
Db	361	GTAATGCATTTCTCTTTAAACCGCGCATGATATGATGATATTTTGAAGTATTTGT	420
QY	1006	GATTCAATGAATCATCATGATGATGATCCCGTACCAAGTTGGCGAAGAAATTTAGCTAGCT	1065
Db	421	GATTCAATGAATCATCATGATGATGATCCCGTACCAAGTTGGCGAAGAAATTTAGCTAGCT	480
QY	1066	ACCGCGGAATTAAGAGATTATTTCAGTTATTCGAACCGGAATTAATTAAGCATCTGCTAGT	1125
Db	481	ACCGCGGAATTAAGAGATTATTTCAGTTATTCGAACCGGAATTAATTAAGCATCTGCTAGT	540
QY	1126	AGTGGCACCATTAATATCCATGATTAATTCATTAATCTCATGAGATAAAAATTTATATGCT	1185
Db	541	AGTGGCACCATTAATATCCATGATTAATTCATTAATCTCATGAGATAAAAATTTATATGCT	600
QY	1186	TATACAGATGAAGAGATTTTTAAAGCCAGCGAGAGTACAAAATTCGAGAAAAATGGCT	1245
Db	601	TATACAGATGAAGAGATTTTTAAAGCCAGCGAGAGTACAAAATTCGAGAAAAATGGCT	660
QY	1246	CAAAACCCATTCAGTGATGATGAGGAGGAGAAAAAATAGTCTGCATAAAGGACTTTCTT	1305
Db	661	CAAAACCCATTCAGTGATGATGAGGAGGAGAAAAAATAGTCTGCATAAAGGACTTTCTT	720
QY	1306	GGAATGTGAATTAAGAAGAACCGGGCGTTGGTATATCGAAAAACTCATCTCTATTAAT	1365
Db	721	GGAATGTGAATTAAGAAGAACCGGGCGTTGGTATATCGAAAAACTCATCTCTATTAAT	780
QY	1366	AAAGATTAATTAAGATTAATCTACTCTTTCGACACACCTCTCGGATTAAGTCCAGGCGCTC	1425
Db	781	AAAGATTAATTAAGATTAATCTACTCTTTCGACACACCTCTCGGATTAAGTCCAGGCGCTC	840
QY	1426	TACGACTTGCTTAGCCAAAAACAACCTAGCTGTCTGATTAATTCATCAAGTTTATTTCA	1485

D <sub>b</sub>	841	AACGACTTGGTTACGCCAAAAACACATCAGCTGCTGATATATTACATACGTTTAAATTC	900
Q <sub>Y</sub>	1486	GCTATTGAGCAGCTGAACCGTTTCATTACAGAAATATGATTCAGTGAACGTCCTCTA	1545
D <sub>b</sub>	901	GCTATTGAGCAGCTGAACCGTTTCATTACAGAAATATGATTCAGTGAACGTCCTCTA	960
Q <sub>Y</sub>	1546	GATGACACGCTCTGGTAAATGA	1566
D <sub>b</sub>	961	GATGACACGCTCTGGTAAATGA	981

RESULT	12				
YEP_LCRGVHP	LOCUS	YEP_LCRGVHP	2201 bp	DNA	BCT 26-APR-1993
DEFINITION		Yersinia pseudotuberculosis V-antigen (LcrG, LcrV, LcrH) genes, complete cds.			
ACCESSION		M57893			
VERSION		g155456			
KEYWORDS		M57893.1	GI:155456		
SOURCE		V-antigen.			
ORGANISM		Yersinia pseudotuberculosis (strain YPIII (PIBI) DNA.			
REFERENCE		Yersinia.			
AUTHORS		1 (bases 1 to 2201)			
TITLE		Bergman, T., Hakansson, S., Forsberg, A., Norlander, L., Maccellaro, A., Baeckman, A., Boelin, I. and Wolf-Watz, H.			
JOURNAL		Analysis of the V antigen lcrGVH-yopBD Operon of Yersinia pseudotuberculosis: Evidence for a regulatory role of LcrH and LcrV			
MEDLINE		J Bacteriol. 173, 1607-1616 (1991)			
FEATURES		91154114			
source		Location/Qualifiers			
		1..2201			
		/organism="Yersinia pseudotuberculosis"			
		/strain="YPIII(PIBI)"			
		/db_xref="taxon:633"			
		219..224			
		/gene="lcr operon"			
		219..258			
		/gene="lcr operon"			
		241..246			
		/gene="lcr operon"			
		252..258			
		/gene="lcr operon"			
		264..551			
		/gene="lcrG"			
		264..551			
		/gene="lcrG"			
		/gene="lcrG"			
		/codon_start=-1			
		/transl_table=11			
		/evidence=experimental			
		/product="V-antigen"			
		/protein_id="AA27644.1"			
		/db_xref="PID:g155457"			
		/db_xref="GI:155457"			
		/translation="MKSHPEDEYDKTLKQAELEIADSDHRAKLQEMCADIGLTPEAV			
		MKIFAGSADIEIKPAERELDELIRQBERQPPQPRPYDOKRRKPTMMRGQII"			
		553..1533			
		/gene="lcrV"			
		553..1533			
		/gene="lcrV"			
		/codon_start=-1			
		/transl_table=11			
		/evidence=experimental			
		/product="V-antigen"			
		/protein_id="AA27645.1"			
		/db_xref="PID:g155458"			
		/db_xref="GI:155458"			
		/translation="MIRAYRONPOHFLEDEKYRVOLTGHSVLEELVQIYKDKNI			
		DISIKDPRKDSVPANRYVTDDIELKLTILAFLEPDALIKGSHYDNOIQNKRYK			
		EFLESSNTQWELRAFMVAVTHFSLIADRIDDDILKVIYDSMNHGDARSLRLELAE			
		TALTKYVIOAEINIKHSSGGTINIHDKSINIMDKNLXGYTDEEIFPKSAEYKILK			



TITLE ViG, a Yersinia enterocolitica lipoprotein involved in Ca2+ dependency, is related to exsb of Pseudomonas aeruginosa

JOURNAL J. Bacteriol. 177 (15), 4230-4237 (1995)

MEDLINE 95362644

REFERENCE 8 (bases 53135 to 53497)

AUTHORS Stainer I., Iriarte M. and Cornelis G.R.

TITLE yscM1 and yscM2, two Yersinia enterocolitica proteins causing downregulation of yop transcription

JOURNAL Mol. Microbiol. 26 (4), 833-843 (1997)

MEDLINE 98087284

REFERENCE 9 (bases 18681 to 18968)

AUTHORS Boyd A.P., Sory M.P., Iriarte M. and Cornelis G.R.

TITLE Heparin interferes with translocation of Yop proteins into HeLa cells and binds to LcrG, a regulatory component of the Yersinia yop apparatus

JOURNAL Mol. Microbiol. 27 (2), 425-436 (1998)

MEDLINE 98143428

REFERENCE 10 (bases 22612 to 22890)

AUTHORS Iriarte M., Sory M.P., Boland A., Boyd A.P., Mills S.D., Lambertmont, I. and Cornelis G.R.

TITLE Lambermont, I. and Cornelis G.R.

JOURNAL TyAa, a protein involved in control of Yop release and in translocation of Yersinia Yop effectors

EMBO J. 17 (7), 1907-1918 (1998)

MEDLINE 98190073

REFERENCE 11 (bases 1 to 69673)

AUTHORS Iriarte M., Lambertmont, I., Kerbourch C. and Cornelis G.R.

TITLE Detailed genetic map of the pYve227 plasmid of Yersinia enterocolitica serotype O:9

JOURNAL Unpublished

REFERENCE 12 (bases 1 to 69673)

AUTHORS Iriarte M., Lambertmont, I., Kerbourch C. and Cornelis G.R.

TITLE Direct Submission

JOURNAL Submitted (30-OCT-1998) Microbial Pathogenesis Unit, Christian de Duve Institute of Cellular Pathology and Faculte de Medecine, Universite Catholique de Louvain, Av. Hippocrate 74.49, Brussels 1200, Belgium

COMMENT On May 18, 1999, this sequence version replaced gi:3411136 gi:3411135 gi:3411139 gi:3411137 gi:3603360 gi:1524380 gi:155536 gi:3603357 gi:2738986 gi:2944089 gi:3411138 gi:3108216 gi:437201 gi:497215 gi:885971 gi:155526 gi:483322 gi:48593 gi:2735421.

FEATURES

SOURCE Location/Qualifiers

1. .69673

/organism="Yersinia enterocolitica"

/strain="W22703"

/serotype="O:9"

/db\_xref="taxon:630"

/plasmid="pYve227"

53. .203

complement(889. .1434)

rep\_origin

CDS

/note="ORF181"

/codon\_start=1

/transl\_table=11

/product="unknown"

/protein\_id="AADI6801.1"

/db\_xref="PID:94324324"

/db\_xref="GI:4324324"

/translation="MRVVIYOGYGIFFKKNGLIDSMPTQTCLGLYALSDOHYLLCA HFDPAFGQONLQDIKRIELINGIMSLRATVFGGDKOSYLRCSAPSSHIGMEIYN FMSHGALAOYSNOYSGVPETEFHFHQKCGIITKGNPNDFLGSGSQANNAKKORI KLRSEVSIYHAKMLDISRLY"

1616. .2083

/note="ORF155"

/codon\_start=1

/transl\_table=11

/product="unknown"

/protein\_id="AADI6802.1"

/db\_xref="PID:94324325"

/db\_xref="GI:4324325"

/translation="MINTFTTELLQKIASHFGLDKLSDDEYGLCELLINDRVIMLRA DEINRFLPLGIPFGSEARSTASOLFFCYSTALNKDGPFCAMSEELGLIAFKHL SIELGVNVSKEIANFYDNLISVSLPMTQDELPHHQSOSTOSTOSYVAKG"

2091. .4280

/gene="yopO"

/translation="MFIKDAYNMRALCTALEQSPADTIIINTSEKNNSYCCATPAHLR

CDS

2091. .4280

/gene="yopO"

/codon\_start=1

/transl\_table=11

/product="protein kinase yopO"

/protein\_id="AADI6803.1"

/db\_xref="PID:94324326"

/db\_xref="GI:4324326"

/translation="MKIMGTPPSISLAKAHERISOHMONPVELNIGKRRYRIIDNQ VLRLNPHSGSELEFREGVKITFSCKMFNSIARNLTEFLHAQKTTSOELSDIPNAL NLFAPKOTLEPLMGKPKPLSGAPDLGMRVAETDPAEGESHISIIETDNORLGR IERSIAGHLEPAELAYAKKHITKTAGKPNLANVGMNVVPGNRKEBALIENDVDGR CSDFLRSLADSMKQKGINSEAYWGTTFIAHRLDVTNNHAKAGTYANDIKPGVAYD RASEPVIYDGLHSRSGEQPKFTSEFKAPELGVDNGASEKSDVEFLVSTLHGIE GFENDPEIKPNQGLRFTYSEPAHYMDNGYPIHPGIAVETAYTRFTIDILGVSAD RPSDENARLHEFLSDGT IDEESAKOILKDTLGTGMSPLDVRRTPEKRLSEDLR THLSAATKOLDMGVSLDDTMTLVDLKAEREGVNDKDKSNLSILKTVSYTEY VKGREDTKSSSAEVSPIYRNSFMELSTAEPSIORLOKHDOOTHSFSDIGSLVRHKL ETLLEVLVYTSPOGQPYSSSEYFSFLNLAEAKVTLSOOLDTLQOQOESAKRQSLTIN RSGSMADVARQSLQRFDSITRVPVAFGEQYETLIRQMAAHAATITLOEVESEFTDMRN FTADSIPLLRIRGRSSLIDELHYOREKRLFTTIAERLNLREEMX"

4485. .4760

/note="ORP91a: unknown function"

/codon\_start=1

/transl\_table=11

/product="unknown"

/protein\_id="AADI6804.1"

/db\_xref="PID:94324327"

/db\_xref="GI:4324327"

/translation="MLSIRKRECPCTWLNGLPVCYGELAINGKPYGINNSALIR IMNLGVLRVLPVCSGFTHORLNLVLOGTYPEYKIHFAVNSLNKDK"

4761. .5627

/gene="yopP"

4761. .5627

/gene="yopP"

/codon\_start=1

/transl\_table=11

/product="Yop effector YopP"

/protein\_id="AADI6805.1"

/db\_xref="PID:94324328"

/db\_xref="GI:4324328"

/translation="MISPIQSINSEFGLSEKETRSLSNBEKNIITQLENDIADGSM FHKYISRLDIEVMPALVYQANNKIPEKMLNVPVSPDLSIEKNVINGGSSFTIN MGEGGIHFVSVDYKHNGKTSLLIFEPVNFMSGPALIASTKTAIERYOLPDPCHFM VEMDIORSSSECGIFSLAKNLKTEYEDSLIKIHEDIKIGLSISEMPLDHNKADPL PVTEFKHOGKRLNEXYINTNPGQVGVVANKNKTIFNRFDNKSIIDGKELSVVK KRIAEYKTLKV"

complement(6151. .6900)

/gene="y1pA"

complement(6151. .6900)

/gene="y1pA"

/codon\_start=1

/transl\_table=11

/product="lipoprotein y1pA"

/protein\_id="AADI6806.1"

/db\_xref="PID:94324329"

/db\_xref="GI:4324329"

/translation="MEDDMKMMLLIATVASSVLVSGGAMSTAIKKRNLEVKTO MSETIWEPSQKRVYVIOIKRNSDKNMLGLAPKTRKAVOKGAVYVSSPDAHWIOA NVLAKDKMDLREAGCFISOGTQGAALTAALGAGITGNSNSAGASLGVGLAAGLVGW ADAVVEDINTMTVDVOISEKTDPILOTDNVAALAKQSTSGYKVTSTOTGKHQYOTR VVSSANKVNLKFEEOQPVLEDOIAKSIANIL"

complement(7370. .7918)

/gene="yopO"

complement(7370. .7918)

/gene="yopO"

/codon\_start=1

/transl\_table=11

/product="YopO"

/protein\_id="AADI6807.1"

/db\_xref="PID:94324330"

/db\_xref="GI:4324330"

/translation="MFIKDAYNMRALCTALEQSPADTIIINTSEKNNSYCCATPAHLR



gene  
TDVCSLVNRVIGIEPLKSGSIISTLEELMOAVGIIILYEMQVSDIDTFNFKLPNNSD  
FGLVFSVLDGDIYITGKDKSGNIEIYDPKNSDLIENDIKKLYDENRFRICML  
IISKESELEDSRESCDOKIMG"  
8439..9407

Query Match 59.2%; Score 927.4; DB 2; Length 69673;

Best Local Similarity 97.3%; Pred. No. 7.9e-193;

Matches 943; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

```
QY 586 ATGATTAGAGCCAGACAAAACCCACAACATTTATTGAGATCTAGAAAAGTTAGG 645
    |||||||
DB 18679 ATGATTAGAGCCAGACAAAACCCACAACATTTATTGAGATCTAGAAAAGTTAGG 18620
QY 646 GTGGAACAACCTTACTGGTCATGGTCTTCAGTTTAGAAGAAATGGTTCAGTTAGTACAA 705
    |||||||
DB 18619 GTGGAACAACCTTACTGGTCATGGTCTTCAGTTTAGAAGAAATGGTTCAGTTAGTACAA 18560
QY 706 GATAAAATATAGATATTTCCATTAAATATGATCCAGAAAAGATTGGAGTTTGGC 765
    |||||||
DB 18559 GATAAAATATAGATATTTCCATTAAATATGATCCAGAAAAGATTGGAGTTTGGC 18500
QY 766 AATAGTATTAATCTGATCATATGCAATTCGCAAGAAATCCTACTTATTTCTACCC 825
    |||||||
DB 18499 GATAGGTAATTAATCTGATCATATGCAATTCGCAAGAAATCCTACTTATTTCTACCC 18440
QY 826 GAGGATACCACTTCTTAAAGCGGCTCATTAATGACACCAACTGCAAAATGGATCAAGCGA 885
    |||||||
DB 18439 GAGGATACCACTTCTTAAAGCGGCTCATTAATGACACCAACTGCAAAATGGATCAAGCGA 18380
QY 886 GTAAAAGAGTCTCTTGAATCATGCGCGAATACCAATGGGAATGGCGGCTTCATGGCA 945
    |||||||
DB 18379 GTAAAAGAGTCTCTTGAATCATGCGCGAATACCAATGGGAATGGCGGCTTCATGGCA 18320
QY 946 GTAAATGCAATTTCTTTAAACCGCCGATCGATGATGATGATATTTTGAAGTGAATGTT 1005
    |||||||
DB 18319 GTAAATGCAATTTCTTTAAACCGCCGATCGATGATGATGATATTTTGAAGTGAATGTT 18260
QY 1006 GATTCATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1065
    |||||||
DB 18259 GATTCATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 18200
QY 1066 ACCGCCGAATTAAGATTTATTCAGTTATTAAGCGGAAATTAAGATCTGTCTAGT 1125
    |||||||
DB 18199 ACCGCCGAATTAAGATTTATTCAGTTATTAAGCGGAAATTAAGATCTGTCTAGT 18140
QY 1126 AATGGACCATTAATATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1185
    |||||||
DB 18139 AATGGACCATTAATATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 18080
QY 1186 TATACAGATGAAGAGATTTTAAAGCGAGGAGACGAGTAAATTCGAGAAATAGCCCT 1245
    |||||||
DB 18079 TATACAGATGAAGAGATTTTAAAGCGAGGAGACGAGTAAATTCGAGAAATAGCCCT 18020
QY 1246 CAACACCACTTCAAGTGGAGCGAGAAAAAATAGTCTCGATTAAGAGACTTCTT 1305
    |||||||
DB 18019 CAACACCACTTCAAGTGGAGCGAGAAAAAATAGTCTCGATTAAGAGACTTCTT 17960
QY 1306 GGAAGTGAAGATTAAGAAACCGGGGCGTTGGTAAATCTGAAAAACATCTTATTAAT 1365
    |||||||
DB 17959 GGAAGTGAAGATTAAGAAACCGGGGCGTTGGTAAATCTGAAAAACATCTTATTAAT 17900
QY 1366 AAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1425
    |||||||
DB 17899 AAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 17840
QY 1426 AAGGACTTGGTATGACCAAAAAACCACTAGCTCTGATTAATTAATTAATTAATTAATTA 1485
    |||||||
DB 17839 AAGGACTTGGTATGACCAAAAAACCACTAGCTCTGATTAATTAATTAATTAATTAATTA 17780
QY 1486 GCTATTGAAGACTGAACGTTTCATTCAGAAATATGATTCAGTATGATGATGATGATGATGAT 1545
    |||||||
DB 17779 GCTATTGAAGACTGAACGTTTCATTCAGAAATATGATTCAGTATGATGATGATGATGATGAT 17720
```

QY 1546 GATGACACG 1554  
|||||  
DB 17719 GATGACACG 17711

RESULT 14  
YE96PVANT YE96PVANT 975 bp DNA BCT 13-MAR-1997  
LOCUS Y.enterocolitica V antigen gene, strain Y-96-P.  
DEFINITION  
ACCESSION X96801  
NID 91405830  
VERSION X96801.1 GI:1405830  
KEYWORDS V antigen.  
SOURCE Yersinia enterocolitica.  
ORGANISM Yersinia enterocolitica  
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

REFERENCE  
AUTHORS 1 (bases 1 to 975)  
Roggenkamp,A., Geiger,A.M., Leitritz,L., Kessler,A. and  
Heesemann,J.  
TITLE Passive immunity to infection with Yersinia spp. mediated by  
anti-recombinant V antigen is dependent on polymorphism of V  
antigen  
JOURNAL Infect. Immun. 65 (2), 446-451 (1997)  
MEDLINE 97162308  
REFERENCE 2 (bases 1 to 975)  
Roggenkamp,A.  
AUTHORS Direct Submission  
TITLE Submitted (22-MAR-1996) A. Roggenkamp, Hygiene and Microbiology,  
University of Wuerzburg, Josef-Schneider Strasse 2, 97080  
JOURNAL Wuerzburg, FRG

## FEATURES

source 1..975  
location/Qualifiers  
/organism="Yersinia enterocolitica"  
/db\_xref="taxon:630"  
1..975  
source /organism="Yersinia enterocolitica (type 0:9)"  
/strain="Y-96-P"  
/db\_xref="taxon:34055"  
1..975  
CDS /codon\_start=1  
/transl\_table=11  
/product="V antigen"  
/protein\_id="CA65593.1"  
/db\_xref="PID:e238988"  
/db\_xref="PID:g1405831"  
/db\_xref="GI:1405831"  
/db\_xref="SPTREMBL:O56882"  
/translation="MIRAYEONPOHFIDLEKVRVDOLTGHSSVLEELVOYVDKKI  
DISIKYDOKNSEVAFADRVITDDIDELKILAIPLPDAIILKGGHYDNOLOGIKRVK  
EFLESSPTOMELRAFVAVMEFSLTADRIDDDILKYVDSMNHHDANGKLEELAEI  
TAELKIVSVIOAEINKHLNSDITNIHDKSINLMDKNLYGTDEEIFKASAEYKILEK  
MPOITIKGEETEKIKVSIKPLESENKRTGALGNLKDYSYKNKNEELSHLGTICSDK  
SRPLNDIVSOYKTOLOSDITSRFNSAIEALNRFIQRKYDSVMRLDDTR"

BASE COUNT 335 a 177 c 194 g 269 t  
ORIGIN

Query Match 58.5%; Score 916.2; DB 1; Length 975;

Best Local Similarity 96.6%; Pred. No. 2.3e-190;

Matches 936; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

```
QY 586 ATGATTAGAGCCAGACAAAACCCACAACATTTATTGAGATCTAGAAAAGTTAGG 645
    |||||||
DB 1 ATGATTAGAGCCAGACAAAACCCACAACATTTATTGAGATCTAGAAAAGTTAGG 60
QY 646 GTGGAACAACCTTACTGGTCATGGTCTTCAGTTTAGAAGAAATGGTTCAGTTAGTACAA 705
    |||||||
DB 61 GTGGAACAACCTTACTGGTCATGGTCTTCAGTTTAGAAGAAATGGTTCAGTTAGTACAA 120
QY 706 GATAAAATATAGATATTTCCATTAAATATGATCCAGAAAAGATTGGAGTTTGGC 765
    |||||||
DB 121 GATAAAATATAGATATTTCCATTAAATATGATCCAGAAAAGATTGGAGTTTGGC 180
```

QY	766	AATAGAGTAATCTACTGATGATGCAATTCGCTCAGAGAAATCCTAGCTTATTTCTACCC	825
Db	181	GATAGGGTAATTAATCTGATGATGCAATTAATCTCAGAGAAATCCTGCTTATTTCTACCC	240
QY	826	GAGATACCACTTCTTAAGGCGGTCATATGACAAACCACTGGCAAAATGGCATCAACGA	885
Db	241	GAGATGCCCTTCTTAAGGCGGTCATATGACAAACCACTGGCAAAATGGCATCAACGA	300
QY	886	GTAAGAGTCTCTTGAATCATCGCCGAATACACAAATGGGAATTCGGCGGCTTCATGGCA	945
Db	301	GTAAGAGTCTCTTGAATCATCGCCGAATACACAAATGGGAATTCGGCGGCTTCATGGCA	360
QY	946	GTAATGATTTCTTCAACCGCGATCGTATCGATGATATTTTGAAGATGATTGT	1005
Db	361	GTAATGATTTCTTCAACCGCGATCGTATCGATGATATTTTGAAGATGATTGT	420
QY	1006	GATTCAATGATCATCTGGGATGCGGATGACAGAGTGGGTGAAGAAATTAGCGACCTT	1065
Db	421	GATTCAATGATCATCTGGGATGCGGATGACAGAGTGGGTGAAGAAATTAGCGACCTT	480
QY	1066	ACCGCGCAATTAAGATTTATTCAGTTATTCAGCCGAATTAATAAGCATCTGTAGT	1125
Db	481	ACCGCGCAATTAAGATTTATTCAGTTATTCAGCCGAATTAATAAGCATCTGTAGT	540
QY	1126	AGTGGCACCAATATATCCATGATTAATTCATTATCTCATGGATTAATAATTTATATGT	1185
Db	541	AGTGGCACCAATATATCCATGATTAATTCATTATCTCATGGATTAATAATTTATATGT	600
QY	1186	TATACAGATTAAGAGATTTTAAAGCCAGCGCAGATACAAATTCGAGAAATATGCT	1245
Db	601	TATACAGATTAAGAGATTTTAAAGCCAGCGCAGATACAAATTCGAGAAATATGCT	660
QY	1246	CAACACCACTTCAAGTGTGATGGAGCGAGAGAAAAAATAGTCTCGATTAAGACTTTCTT	1305
Db	661	CAACACCACTTCAAGTGTGATGGAGCGAGAGAAAAAATAGTCTCGATTAAGACTTTCTT	720
QY	1306	GGAAGTGAAGTAATAAGAACCGGGGCGTTGGTATATCGAAAACTATCTCTTAAAT	1365
Db	721	GGAAGTGAAGTAATAAGAACCGGGGCGTTGGTATATCGAAAACTATCTCTTAAAT	780
QY	1366	AAAGATATATGATATATCTCACTTGTCCACCACTGCTCGGATTAAGTCCAGGCGCTC	1425
Db	781	AAAGATATATGATATATCTCACTTGTCCACCACTGCTCGGATTAAGTCCAGGCGCTC	840
QY	1426	AACGACTTGTAGCCAAAAACAACACACAGCTGTGTGATATATACATCAAGCTTTAATCA	1485
Db	841	AACGACTTGTAGCCAAAAACAACACACAGCTGTGTGATATATACATCAAGCTTTAATCA	900
QY	1486	GCTATGTAACCACTGAACCGTTTCATTCAGAAATATATTCAGTATGCAACGCTGCTTA	1545
Db	901	GCTATGTAACCACTGAACCGTTTCATTCAGAAATATATTCAGTATGCAACGCTGCTTA	960
QY	1546	GATGACACG 1554	
Db	961	GATGACACG 969	
RESULT	15		
LOCUS	YE108VANT	975 bp	DNA
DEFINITION	X. enterocolitica V antigen gene, strain Y-108-P.		
ACCESSION	X96799		
NID	91405822		
VERSION	X96799.1	GI:1405822	
KEYWORDS	V antigen.		
SOURCE	Yersinia enterocolitica.		
ORGANISM	Yersinia enterocolitica.		
REFERENCE	1 (bases 1 to 975)		
AUTHORS	Roggenkamp, A., Geiger, A.M., Leitritz, L., Kessler, A. and Heesemann, U.		

TITLE	Passive immunity to infection with Yersinia spp. mediated by anti-recombinant V antigen is dependent on polymorphism of V antigen				
JOURNAL	Infect. Immun. 65 (2), 446-451 (1997)				
MEDLINE	97162308				
REFERENCE	2 (bases 1 to 975)				
AUTHORS	Roggenkamp, A.				
TITLE	Direct Submission				
JOURNAL	Submitted (22-MAR-1996) A. Roggenkamp, Hygiene and Microbiology, University of Wuerzburg, Josef-Schneider Strasse 2, 97080 Wuerzburg, FRG				
FEATURES	Location/Qualifiers				
source	1..975				
	/organism="Yersinia enterocolitica"				
	/sub_species="type 0:3"				
	/strain="Y-108-P"				
	/db_xref="taxon:630"				
CDS	1..975				
	/codon_start=1				
	/transl_table=11				
	/product="V antigen"				
	/protein_id="CAA65591.1"				
	/db_xref="PID:e23985"				
	/db_xref="PID:g1405823"				
	/db_xref="GI:1405823"				
	/db_xref="SPTREMBL:Q56846"				
	/translation="MIRAEQNPQHFIDLEKRVBOLTHGSSVLEELVOLVODKKI DISIKYDPKDESEFANRYITDIELKILAFLEPEDAILKGHYNDLOJGIRKRV ELESPPNOMELRAFMVHMFSLNADRIDDLILVIYDSNMHGDARSKRELAEL TAEIKYSVIOAKIEINKHLSNSDPTINHKSTINKYKNGYNEIEIFASAEYKILEK MSQTIKBEETKEKIVSINKNEFESERKRIKAGNDKSYKNDNNEHJSTGTSCK SRPLNDLVQKTTQSLDITSRFNSAIEALNRFIQKIDYSVMQRLDDTR"				
BASE COUNT	341 a	173 c	190 g	271 t	
ORIGIN					
Query Match	58.4%;	Score 914.6;	DB 1;	Length 975;	
Best Local Similarity	96.5%;	Pred. No. 5.2e-190;			
Matches 935; Conservative	0;	Mismatches 34;	Indels 0;	Gaps 0;	
QY	586	ATGATTAAGACCTACGACCAAAACCAACCAATTTATGAGATCTAGAAAAAGTTAGG	645		
Db	1	ATGATTAAGACCTACGACCAAAACCAACCAATTTATGAGATCTAGAAAAAGTTAGG	60		
QY	646	GTGAAACAATTAAGTGTGATGGTCTTTCAGTTTGAAGAATTTGGTTCAAGTATGCAAA	705		
Db	61	GTGAAACAATTAAGTGTGATGGTCTTTCAGTTTGAAGAATTTGGTTCAAGTATGCAAA	120		
QY	706	GATAAAAATATGATATTTCCATTAAATATGATCCAGAAAAGATTGGAGCTTTTGGC	765		
Db	121	GATAAAAATATGATATTTCCATTAAATATGATCCAGAAAAGATTGGAGCTTTTGGC	180		
QY	766	AATAGATTAATTAAGTATGATGATGATGCTCAAGAAATCCTAGCTTATTTCTACCC	825		
Db	181	AATAGGTTAATTAAGTATGATGATGATGCTCAAGAAATCCTAGCTTATTTCTACCC	240		
QY	826	GAGATACCACTTCTTAAGGCGGTCATATGACAAACCACTGGCAAAATGGCATCAACGA	885		
Db	241	GAGATGCCCTTCTTAAGGCGGTCATATGACAAACCACTGGCAAAATGGCATCAACGA	300		
QY	886	GTAAGAGTCTCTTGAATCATCGCCGAATACACAAATGGGAATTCGGCGGCTTCATGGCA	945		
Db	301	GTAAGAGTCTCTTGAATCATCGCCGAATACACAAATGGGAATTCGGCGGCTTCATGGCA	360		
QY	946	GTAATGATTTCTTCAACCGCGATCGTATCGATGATATTTTGAAGATGATTGT	1005		
Db	361	GTAATGATTTCTTCAACCGCGATCGTATCGATGATATTTTGAAGATGATTGT	420		
QY	1006	GATTCAATGATCATCTGGGATGCGGATGACAGATTCGTTGAAGAAATTAGCATCTGTAGT	1065		
Db	421	GATTCAATGATCATCTGGGATGCGGATGACAGATTCGTTGAAGAAATTAGCATCTGTAGT	480		
QY	1066	ACCGCGCAATTAAGATTTATTCAGTTATTCAGCCGAATTAATAAGCATCTGTAGT	1125		

Db	481	ACCCGCAATTAATAAATTATTCAGTTATTCAGGCTGAATTAATAAGACTCTGCAAT	540
Oy	1126	AGTGGCACCATTAATATCCATCGATTAATTCATTAACTCTATGATATAAATTTATATGT	1185
Db	541	AGTACACCCATTAATATCCATGATTAATCATTAATCTCATGATAAATAATTTATATGT	600
Oy	1186	TATACAGTGAAGATTTTTTAAAGCAGCGCAGATACAAATTCGAGAAATGGCT	1245
Db	601	TATCAAAATGAAGATTTTTTAAAGCCAGCGCAGACTACAAATTTCTGAGAAATGGCT	660
Oy	1246	CAAACCCACATTCAGTGGATGGAGCGAGAAAAAATATAGTCTCATTAAGACATTTCTT	1305
Db	661	CAAACCCACATTAAGAGAGGTGAGACCGAAAAAATATAGTCTCATTAAGACATTTCTT	720
Oy	1306	GGAAGTGAGATTAAGAACCAGGGGCGTTGGTATCTGATAAAACTCATCTTTATAT	1365
Db	721	GAAAGTGAGATTAAGAAMCCGGGGCGTTGGGTAACTGAAGATTCATCTTTATAT	780
Oy	1366	AAAGTATATATGAATATCTCACTTGTGCACACCTGTCGGATTAATGCCAGGCGCTC	1425
Db	781	AAAGTATATATGAATATATCTCACTTGGGACACCTGTCGGATTAATGCCAGGCGCTC	840
Oy	1426	AACGACTTGGTTACCAAAAAACACTGCAGCTCTGATATTACATCAAGTTTATATCA	1485
Db	841	AACGACTTGGTTACCAAAAAACACTGCAGCTCTGATATTACATCAAGTTTATATCA	900
Oy	1486	GCTATTGAGACATGAACGTTTTCAATTGAGAAATATGATTCAAGTATGCAACGTCGCTA	1545
Db	901	GCTATTGAGACATGAACGTTTTCAATTGAGAAATATGATTCAAGTATGCAACGTCGCTA	960
Oy	1546	GATGACAGC 1554	
Db	961	GATGACAGC 969	

```
Search completed: November 27, 1999, 18:43:12
Job time: 1232 sec
```

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 1999, 17:32:44 ; Search time 305.3 Seconds  
(without alignments)  
12015.090 Million cell updates/sec

Title: US-08-699-716A-1  
Perfect score: 1566  
Sequence: 1 ATGGGCCATCATCATCATCA.....ATGACACGTCGTGTAATGCA 1566

Scoring table: IDENTITY\_NOC

Searched: 3032314 segs, 1171202697 residues

Database :

EST: \*  
1: em\_est1: \*  
2: em\_est2: \*  
3: em\_est3: \*  
4: em\_est4: \*  
5: em\_est5: \*  
6: em\_est6: \*  
7: em\_est7: \*  
8: em\_est8: \*  
9: em\_est9: \*  
10: em\_est10: \*  
11: em\_est11: \*  
12: em\_est12: \*  
13: em\_est13: \*  
14: em\_est14: \*  
15: em\_est15: \*  
16: em\_est16: \*  
17: em\_est17: \*  
18: em\_est18: \*  
19: em\_est19: \*  
20: em\_est20: \*  
21: em\_est21: \*  
22: em\_est22: \*  
23: em\_est23: \*  
24: em\_est24: \*  
25: em\_est25: \*  
26: em\_est26: \*  
27: em\_est27: \*  
28: em\_est28: \*  
29: em\_est29: \*  
30: em\_est30: \*  
31: em\_est31: \*  
32: em\_est32: \*  
33: em\_est33: \*  
34: em\_est34: \*  
35: em\_est35: \*  
36: em\_est36: \*  
37: em\_est37: \*  
38: em\_est38: \*  
39: em\_est39: \*  
40: em\_est40: \*  
41: em\_est41: \*  
42: em\_est42: \*  
43: em\_est43: \*  
44: em\_est44: \*  
45: em\_est45: \*  
46: em\_est46: \*  
47: em\_est47: \*  
48: em\_est48: \*  
49: em\_est49: \*  
50: em\_est50: \*  
51: em\_est51: \*  
52: em\_est52: \*  
53: em\_est53: \*

54: em\_est22: \*  
55: em\_est23: \*  
56: em\_est24: \*  
57: em\_est25: \*  
58: em\_est26: \*  
59: em\_est27: \*  
60: em\_est28: \*  
61: em\_est29: \*  
62: em\_est30: \*  
63: em\_est31: \*  
64: em\_est32: \*  
65: em\_est33: \*  
66: em\_est34: \*  
67: em\_est35: \*  
68: em\_est36: \*  
69: em\_est37: \*  
70: em\_est38: \*  
71: em\_est39: \*  
72: em\_est40: \*  
73: em\_est41: \*  
74: em\_est42: \*  
75: em\_est43: \*  
76: em\_est44: \*  
77: em\_est45: \*  
78: em\_est46: \*  
79: em\_est47: \*  
80: em\_est48: \*  
81: em\_est49: \*  
82: em\_est50: \*  
83: em\_est51: \*  
84: em\_est52: \*  
85: em\_est53: \*  
86: em\_est54: \*  
87: em\_est55: \*  
88: em\_est56: \*  
89: em\_est57: \*  
90: em\_est58: \*  
91: em\_est59: \*  
92: em\_est60: \*  
93: em\_est61: \*  
94: em\_est62: \*  
95: em\_est63: \*  
96: em\_est64: \*  
97: em\_est65: \*  
98: em\_est66: \*  
99: em\_est67: \*  
100: em\_est68: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	41.8	2.7	579	44	AU038989	AU038989 AU038989
2	41	2.6	444	24	H78200	H78200 yu84f10.r1
3	40.8	2.6	561	40	C91275	C91275 C91275 Dict
4	40.6	2.6	394	25	N97614	N97614 1054C3 czap
5	39	2.5	510	39	C84881	C84881 C84881 Dict
6	38.6	2.5	1107	31	C22922	C22922 C22922 Dict
7	38.6	2.5	526	63	AM004525	AM004525 701932008
8	38.6	2.5	489	64	AL121551	AL121551 DKF28762L
9	38	2.4	243	35	AA547782	AA547782 EST188681
10	37.8	2.4	469	62	AU076358	AU076358 AU076358
11	37.4	2.4	640	26	W95009	W95009 zh46d04.r1
12	37.4	2.4	685	50	AU061085	AU061085 AU061085
13	37.2	2.4	506	61	AI808691	AI808691 w57408.x
14	37	2.4	229	35	AA547780	AA547780 EST188689
15	37	2.4	496	44	AU037010	AU037010 AU037010
16	36.8	2.3	382	29	AA143875	AA143875 MCCCE0261.
17	36.8	2.3	249	35	AA547759	AA547759 EST188668
18	36.8	2.3	239	35	AA547760	AA547760 EST188669
19	36.6	2.3	270	35	AA547784	AA547784 EST188693
20	36.4	2.3	588	35	AA550687	AA550687 1924m3 gm
21	36.4	2.3	689	64	AL119439	AL119439 DKF2P761K
22	36.2	2.3	424	20	D29133	D29133 HDNKK157 Hu
23	36.2	2.3	401	37	AA668776	AA668776 aa80b10.s
24	36.2	2.3	422	59	AI775965	AI775965 EST257065
25	36.2	2.3	522	59	AI776377	AI776377 EST257477
26	36	2.3	229	28	AA114611	AA114611 mm04C03.r
27	36	2.3	625	44	AU037754	AU037754 AU037754
28	36	2.3	327	44	AU037911	AU037911 AU037911
29	35.8	2.3	480	22	H09719	H09719 y190b12.s1
30	35.8	2.3	573	39	AA894968	AA894968 vx65C06.r
31	35.6	2.3	421	35	C24612	C24612 C24612 Dict
32	35.6	2.3	384	40	AA978194	AA978194 OP33403.S
33	35.6	2.3	291	44	AU039780	AU039780 AU039780
34	35.4	2.3	400	20	T09814	T09814 0401m3 gmbp
35	35.4	2.3	267	20	T18170	T18170 0770C3 czap
36	35.4	2.3	221	35	AA547786	AA547786 EST188695
37	35.4	2.3	406	40	AA975418	AA975418 cm59g12.s
38	35.4	2.3	433	41	AI027570	AI027570 oy46C06.x
39	35.4	2.3	409	43	AI186323	AI186323 qd25h10.x
40	35.4	2.3	405	44	AI279748	AI279748 qp28a09.x
41	35.4	2.3	566	46	AI421821	AI421821 lf55b04.x
42	35.4	2.3	431	50	AI673292	AI673292 tw69g08.x
43	35.4	2.3	657	63	AI935161	AI935161 wp14d12.x
44	35.4	2.3	569	64	AW052061	AW052061 wx25b04.x
45	35.2	2.2	435	22	R80620	R80620 y196f09.r1

## ALIGNMENTS

RESULT	1
AU038989	
LOCUS	AU038989
DEFINITION	AU038989 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
ACCESSION	AU038989
NID	93985742
VERSION	AU038989.1
KEYWORDS	GI:3985742
SOURCE	ESF
ORGANISM	Dictyostelium discoideum.
REFERENCE	Dictyostelium discoideum.
AUTHORS	Eukaryota; Dictyostelida; Dictyostelium. 1 (bases 1 to 579)
	Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M., Yoshino,R., Mitra,B.N., Pl.M., Sato,T., Takemoto,K., Yasukawa,H., Williams,J., Maeda,M., Takeuchi,T., Ochiai,H. and Tanaka,Y. The Dictyostelium developmental cDNA project: generation and analysis of expressed sequence tags from the first-finger stage of development
JOURNAL	DNA Res. 5 (6), 335-340 (1998)
MEDLINE	99156227
COMMENT	On Jan 19, 1998 this sequence version replaced gi:2150712.

CONTACT: Hideko Urushihara  
Institute of Biological Sciences  
University of Tsukuba  
3-3-10 Ten-noda, Tsukuba, Ibaraki 305, Japan  
Email: d402h@tsukuba.ac.jp  
PROJECT = 'Dictyostelium discoidium cDNA project in Japan'  
Location/Qualifiers  
1..579

/organism="Dicycoscellium discoideum"  
 /strain="AX4"  
 /db\_xref="taxon:44689"  
 /map="2: 21q"  
 /clone="SSM246"  
 /clone.lib="Dictyostelium discoideum SS (H.Urushihara)"  
 /dev\_stage="slug"  
 BASE COUNT 286 a 31 c 52 g 210 t  
 ORIGIN

Query Match	2.7%	Score 41.8	DB 44	Length 579
Best Local Similarity	45.0%	Pred. NO. 1.4		
Matches 157, Conservative	0	Mismatches 192	Indels 0	Gaps 0

QY	970	GATGCTATCATGATATGATATTTTGAAGAGCATTTGATTCATGATCATCATGAT	1029
Dd	71	GGCATATTTTGGATGGAATATTATTAATATCATAGTAATTAATTAATTAATGACAT	130
QY	1030	GCCCGTAGCAATGGCTGAAGAAATTAAGCTGACCTTACCCCGCAATTAAGATTATTC	1088
Dd	131	GATGATCATGATTAATAAATGAAGAAGATGATGATTAATAATTAATTAATTAATATAT	190
QY	1090	GTTATTCAGCCGAATTAATTAAGCATCTGTCTAGTAGTGGCACCATAATATTCATGAT	1149
Dd	191	AATATATATATATATATATATTAATTAATTAATTAATTAATTAATTAATTAATATAT	250
QY	1150	AAATCATTAATCTCATGATTAATAAATTTATATGTTATATCAGATGAGAGATTTTAA	1209
Dd	251	AATATATATATATTAATTAATTAATTAATTAATTAATTAAGAATTTATTAATTTTAAT	310
QY	1210	GCCAGCCAGTAGTCAAAATTTCTGAGAAATCCCTCAACACACCATCTAGCTGGATGG	1268
Dd	311	TGATATTAACATGATATGATTAATCAATCATATTTTCAATTAATAATTCAGAGGTCT	370
QY	1270	AGCAGAGAAAAATATGCTCTGATTAAGGACTTCTTGGAGTGAGATA	1318
Dd	371	ATGAACACAGGTGTTTATTTAAAAATGAATATTAATGAATTAATGAAA	419

RESULT	2
H78200	
LOCUS	
DEFINITION	yH8200. 444 bp mRNA EST 09-NOV-1995
ACCESSION	YH84110.r1 Soares fetal liver spleen 1mFLS Homo sapiens cDNA clone
NID	IMAGE:240523.5 similar to gb U67929 HUMALUL16 Human carcinoma
VERSION	cell-derived Alu RNA transcript, (RNA); gb:M57627 INTERLEUKIN-10
KEYWORDS	PRECURSORS (HUMAN); contains Alu repetitive element; mRNA sequence
SOURCE	H78200 G1056289 H78200.1 GI:1056289 EST.
ORGANISM	human. Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 444) Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
TITLE	The Washu-Merck EST Project
JOURNAL	Unpublished (1995)
COMMENT	On Apr 14, 1993 this sequence version replaced gi:692889.

Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 Insert Size: 1655  
 High quality sequence stops: 342  
 Source: IMAGE Consortium, LNL  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium ([info@image.lnl.gov](mailto:info@image.lnl.gov)) for further information  
 Insert Length: 1655 Std Error: 0.00  
 Seq primer: M13Rpl  
 High quality sequence stop: 342.  
 Location/Qualifiers  
 1..444

```

/organism="Homo sapiens"
/db_xref="GDB:3789656"
/db_xref="taxon:9606"
/clone="IMAGE:240523"
/clone_lib="Soares fetal liver spleen 1NFLS"
/sex="male"
/dev_stage="20 week post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pTV3D (Pharmacia)
with a modified polylinker; site_1: Pac I; Site_2: Eco RI
1st strand cDNA was primed with a Pac I - oligo(dT) primer
13 AACGCGAGAGATTAAATTAAGATCTTTTCTTTTCTTTT 3',
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pTV3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fátima Bonaldo."

```

Query Match	2.6%	Score 41;	DB 24;	Length 444;
Best Local Similarity	60.2%	Pred. No. 2.2;		
Matches 68; Conservative	0;	Mismatches 45;	Indels 0;	Gaps 0;
QY 1081	ATTATTCAGTTATTCACGCCGAATTTATTAAGCATCTGTCAGTAGAGGCACCATTAAT	1140		
Db 193	ATTATTTAGTCATTTCACGAAGTATTTGAATGAGCATCTGCCAGGGCTCGGAATGTTTAT	252		
QY 1141	ATTCATGATTAATTCATTAATCTCATGTGATTAATAATTTATATGTTATACCA	1193		
Db 253	AACACTGCTTATTTATTTTACTTATTTATTTATTTATTTATTTTATTTGAGACACA	305		

	RESULT	3	C91275	561 bp	mRNA	EST	20-APR-1998
	LOCUS		C91275				
	DEFINITION		C91275 Dictyostelium discoideum SS (H.Urushihara)			Dictyostelium	
	ACCESSION		C91275			dicoideum cDNA clone S5J737, mRNA sequence.	
	NID		G3060641				
	VERSION		C91275.1	GI:3060641			
	KEYWORDS		EST,				
	SOURCE		Dictyostelium discoideum,				
	ORGANISM		Eukaryota; Dictyosteliida;			Dictyostelium.	
	REFERENCE		Yoshino,R., Morio,T. and Tanaka,Y.				
	AUTHORS		Developmental cDNA In Dictyostelium discoideum				
	TITLE		Unpublished (1997)				
	JOURNAL		On Jan 19, 1998 this sequence version replaced gi:2152757.				
	COMMENT		Contact: Hideko Urushihara, Institute of Biological Sciences University of Tsukuba 3-3-10 Ten-noda1, Tsukuba, Ibaraki 305, Japan Email: d402hu@esakura.cc.tsukuba.ac.jp.  location/Qualifiers				
	FEATURES		1..561				
	SOURCE		/organism="Dictyostelium discoideum"				
			/strain="AX4"				
			/db_xref="taxon:44689"				
			/map_expr="17q21"				
			/clone="SSJ737"				
			/clone_lib="Dictyostelium discoideum SS (H.Urushihara)"				
			/dev_stage="slug"				
	BASE COUNT		273 a	86 c	56 g	146 t	
	ORIGIN						
OY	Query Match		Best Local Similarity	2.6%:	Score 40.8:	DB 40:	Length 561:
Dd	Matches	54:	Conservative	0:	Mismatches	22:	Indels
OY		66 GCATATGAAAAAATC	81				
Dd		308 CCATTATAAAAGCAAC	323				
RESULT	4						
LOCUS	N97614	394 bp	mRNA	EST	18-NOV-1996		
DEFINITION	I054C3 czaprendz.1, Deoopen Chakrabarti Plasmodium falci-parum CDNA						
ACCESION	cione PF1054c, mRNA sequence.						
NID	N97614						
KEYWORDS	g1674632						
SOURCE	N97614.1 GI:1674632						
ORGANISM	malaria parasite P. falciparum.						
REFERENCE	Plasmodium falci-parum						
AUTHORS	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.						
	Dane,J.B., Annot,D.E., Bourke,P., Chakrabarti,D., Christodoulou,Z., 1 (bases 1 to 394) Hinterberg,K., Holder,A.A., Holt,C.D., Kemp,D., Lanzer,M., Lim,A., Newbold,C., Ravetch,J.V., Reddy,G.R., Rubio,J.J., Schuster,S.M., Su,X.-Z., Thompson,J.K., Vital,F., Williams,T.E. and Werner,E. Current status of the plasmodium falciparum genome project Mol. Biochem. Parasitol. 79, 1-12 (1996) 97001675						
COMMENT	On May 5, 1995 this sequence version replaced gi:798390.						

	Contact: Debopam Chakrabarti Department of Molecular Biology and Microbiology University of Central Florida Orlando, FL 32816-2360 Tel: 407 384 2061 Fax: 407 384 3095 Email: dchakepegasus.cc.ucf.edu Seq primer: T3.
FEATURES	Location/Qualifiers
Source	1..394
	/organism="Plasmodium falciparum"
	/strain="Dd2"
	/db_xref="taxon:5833"
	/map="21"
	/clone="ppf1054C"
	/clone_1lb="czappPd2.1, Debopam Chakrabarti"
	/lab_host="E. coli XL-1 blue"
	/note="Vector: Lambda ZAP II; Site_1: EcoR I; Site_2: Xho I; PolyA+ RNA, from asynchronous blood stage parasites of the Dd2 isolate cultured in vitro. Was reverse transcribed using an oligo dt-Xho I primer. Second strand was prepared using RNase H and DNA polymerase I. EcoR I adapters were ligated to the cDNA, and it was digested with Xho I. Prepared fragments were ligated into EcoR I + Xho I digested lambda ZAP II vector."
BASE COUNT	229 a 36 c 34 g 94 t 1 others
ORIGIN	
Query Match	2.6%; Score 40.6; DB 25; Length 394;
Best Local Similarity	45.5%; Pred. No. 2.7;
Matches 142:	Conservative 0; Mismatches 170; Indels 0; Gaps 0;
OY	1077 AAGAGTTATTACGTTATTCAGCCGAANTTATTAAGCATCTGCTAGTAGCGCACCAT 1136
Db	65 AGATATATCTCTNNATTTTCATAAAAATAAACCACTTCATCAACAATAAAAAACAT 124
OY	1137 AAATATCATGTAAATGCATTAATCTCATGGATGATAAAATTATATNGTTATACAGATGA 1196
Db	125 ATACAAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 184
OY	1197 AGAGATTTTTTAAAGCCAGCGCAGAGTACAAAATTTCTCGAANAATGCCCAACCACCAT 1256
Db	185 AGCTATAGACATGAGAAGAAACACATTTGCTCAAAATAACAACAANAATGTGAATATACATTA 244
OY	1237 TCAGGTGCATGGGACCGAGAAAAAATAGTCTCGATTAAGGACTTTTGGAAGTAGAGAA 1316
Db	245 TATGATATATCATATGATCAACAAATAAAAAATTAACATCAAAATTTTGAAGAAGCAAACA 304
OY	1317 TAAAGAACCGGGGCTGGGTAATCTGAAAACATCATCTTTATTAATTAAGATATATA 1376
Db	305 AAATGTAAATTAACCTATATATTAATTTAAAGATTGGGAAATTAATACTTTTAAAAATGA 364
OY	1377 TGAATTATCTCA 1388
Db	365 CGAAAAATATATA 376
RESULT 5	
LOCUS	C84881 510 bp mRNA EST 12-JUL-1999
DEFINITION	C84881 Dictyostelium discoideum SS (H Urushihara) Dictyostelium
ACCESSION	C84881 dictioideum cDNA clone SSD366, mRNA sequence.
NID	g2865028
VERSION	C84881.1 GI:2865028
KEYWORDS	EST.
SOURCE	Dictyostelium discoideum.
ORGANISM	Dictyostelium discoideum.
REFERENCE	Eukaryota, Dictyostellida: Dictyostelium.
AUTHORS	Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,





**Contact:** Ottenwaelder B  
MPS  
Am Klopferpfütz 18a D-82152 Martinsried, Germany  
This is the 5' sequence of the clone insert  
clone Itom 5. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;  
sequenced by MedicGenomix within the cDNA sequencing consortium of  
the German Genome Project.  
No SI sequence available.  
This clone is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

Query Match	2.5%	Score 38.6	DB 64	Length 489
Best Local Similarity	55.6%	Pred. No. 9.1		
Matches 74	Conservative 0	Mismatches 59	Indels 0	Gaps 0

[illegible]

RESULT	9
AA547782	
LOCUS	243 bp mRNA
DEFINITION	EST188651 B. glabrata adult cerebral ganglia Biophthalmaria glabrata
DEFINITION	CDNA 5' similar to mitochondrial, mRNA sequence.
DEFINITION	AA547782
DEFINITION	28-SEP-1999

REFERENCE (bases 1 to 243)  
AUTHORS Knight, M., Miller, A.N., Geoghagen, N.S.M., Lewis, F.A. and  
Kerlavage, A.R.  
TITLE Expressed sequence tags (ESTs) of *Biomphalaria glabrata*, an  
intermediate snail host of *Schistosoma mansoni*: Use in the  
identification of RFLP markers  
JOURNAL *Malacologia* 39 (1-2), 175-182 (1998)  
COMMENT On Apr 14, 1993 this sequence version replaced gi:562984.

Contact: Kerlavage, AR  
Bioinformatics  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850 USA  
Tel: 3018699056  
Fax: 3018699423  
Email: arkerlavage@igf.org  
Seq primer: M13 Reverse.

```

FEATURES
SOURCE
    Location/Qualifiers
      1. 243
        /organism="Biomphalaria glabrata"
        /strain="BS-90"
        /db_xref="taxon:6526"
        /map="MMU16C3-C4 region; MMU16C3-C4 region"
        /clone_lib="B. glabrata adult cerebral ganglia"
        /tissue_type="cerebral ganglia"
        /note="Organ: brain; Vector: plasmid; Site_1:
        EcoRI; Site_2: XhoI; Total RNA was purified from adult
        cerebral ganglia exposed to S. mansoni miracidia. cDNA
        was constructed using an oligo-dT primer and directionally
        cloned using the Lambda Zap II Vector Kit by Stratagene."
BASE COUNT
    104 a      12 c      26 g      98 t      3 others
ORIGIN

```

Query Match Similarity	2.4%	Score	38	DB	35	Length	243
Best Local Similarity	51.2%	Pred. No.	12				
Matches	86	Conservative	0	Mismatches	82	Indels	0
				Gaps			0

QY	1073	AATTAAGATTTTATAGTATATCAAGCCGAATTAATGAGCATGTGCTAGTATGGCA	1132
Db	56	AATGAAAAAATTAAAGATTAATTTTCCGGAATTTATGTAAGTAATAAAAAATTAAT	115
QY	1133	CCATAATATCCATGATTAATCCATTAATCTCATGATTAATTAATAGTTTACAG	1192
Db	116	AAGCAGTATAGTTAATATATATTTGTTTACAGGAATGAAAAATTTNCATTATAATG	175

QY 1193 ATGAAGACATTTTAAAGCGAGCAGATCAAAATTCTCAGAGAAA 1240  
 || | ||||| | | ||||| || |  
 Db 176 ATAACTGATTTTATAATAAGTATTTAGTCTTAATACAAATTA AAA 223

## RESULT 10

[illegible]

DB	220	ATTATTAAGCTGTATAACAGCCTTCACCTTTGGAAACATTTCAAGGACGATP	170
OY	900	TGAATCATTCGCCGATATACACAATGGGAATTCGGCGCTCAATGCAGATAT	950
DB	280	TAAATCGGGGCGATGCTCTTAAACAGCCTTCAGATATATTTCAAGCAAGTAAAGGAGATCC	221
OY	840	TAAAGCGGCTATATATGACAACAACATCACAATGGCATCAAGCGAGTAAAGAGTTCT	899
DB	340	TGATTTTAAACAGCCTTTTACAATGTTTTTAATTTGATTATACGTATATGCTTATTTGGA	281
OY	780	TGATGATATCGAATTCGTCAGAAGAAAATCTGACTATTTTGTAACCCGGAGATACATTC	839
DB	400	TCTTTTTCACATTTGGTTTTCTTAAACCACACATGTATGTTCCTCAATTAATAAAT	341
OY	720	TATTTCCATTAAATATATGATCCACAGAAAAGATTCGAGGTTTTTGCCAATAGAGTATATAC	779
DB	400	TCTTTTTCACATTTGGTTTTCTTAAACCACACATGTATGTTCCTCAATTAATAAAT	341
OY	240	Best Local Similarity	47.6%; Pred. No. 19;
MATCHES	110;	Conservative	0; Mismatches 121; Indels 0; Gaps 0;
BASE COUNT	192 a	128 c	123 g 193 t 4 others
ORIGIN			
FEATURES	SOURCE		
COMMENT	Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: estewatson.wustl.edu This clone is available royalty-free through LNL ; contact the IMAGE Consortium (infoimage.lnl.gov) for further information. Insert Length: 1189 Std Error: 0.00 Seq primer: mob.REGA+ER High quality sequence stop: 440.		
REFERENCE	AUTHORS		
JOURNAL MEDLINE	TITLE		
GENOME RES.	Generation and analysis of 280,000 human expressed sequence tags		
97044478			
ORGANISM	human.		
SOURCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 640)		
AUTHORS	Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chisoe,S., Dietrich,N., Dubugue,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Laczy,M., Le,M., Le,N., Mardis,E., Moore,B., Morris,M., Parsons,J., Pegeon,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J., Trevaskis,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson.R. and Marra,M.		





GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: November 27, 1999, 18:49:42 ; Search time 43.65 Seconds

(without alignments)  
8975.963 Million cell updates/sec

Title: US-08-699-716A-1

Perfect score: 1566  
Sequence: 1 ATGGGCCATCATCATCATCA.....ATGACAGCTGTGTGAATGA 1566

Scoring table: IDENTITY\_NUC

Searched: 311585 seqs, 125096042 residues

Database : N\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1472	94.0	1530	1 T38249	Y. pestis Fl/V ant
2	1409	90.0	1462	1 T38256	Y. pestis Fl/V ant
3	976.8	62.4	1014	1 T38222	Partial LcrV (V an
4	975.2	62.3	1014	1 T38242	Y. pestis lcrV (V
5	974.8	62.2	1014	1 T38243	Y. pestis lcrV (V
6	973.2	62.1	1014	1 T04223	Partial LcrV (V an
7	517.2	33.0	544	1 V41594	Nucleotide sequenc
8	516.8	33.0	547	1 O92819	Yersinia pestis ca
9	516.8	33.0	544	1 T38248	Y. pestis Fl antige
10	516.8	33.0	544	1 V41595	Nucleotide sequenc
11	510	32.6	474	1 V41600	Nucleotide sequenc
12	452.8	28.9	474	1 V41600	Nucleotide sequenc
13	452.2	28.9	541	1 O92817	Yersinia pestis ca
14	452.2	28.9	541	1 T38244	Y. pestis caFl (Fl
15	451	28.8	542	1 O92818	Yersinia pestis ca
16	448	28.6	450	1 V41601	Nucleotide sequenc
17	447	28.5	447	1 V41609	Nucleotide sequenc
18	402	25.7	576	1 V41598	Nucleotide sequenc
19	339	21.6	513	1 V41599	Nucleotide sequenc
20	68.8	4.4	861	1 O81500	sfy anti-rev seque
21	68.8	4.4	861	1 T45347	Single chain sfy a
22	68.4	4.4	375	1 V54056	DNA fragment 2 of
23	68.4	4.4	375	1 V54057	DNA fragment 3 of
24	68.4	4.4	375	1 V54058	DNA fragment 4 of
25	54	3.4	1546	1 V30575	Clostridium botuli
26	53.4	3.4	1402	1 T29246	Type A neurotoxin
27	53.4	3.4	1278	1 V41263	Chlamydomonas rein
28	53.4	3.4	1469	1 V30591	Clostridium botuli
29	53.4	3.4	1460	1 V30593	Clostridium botuli
30	53.4	3.4	1535	1 V30596	Clostridium botuli
31	53.4	3.4	1526	1 V30580	Clostridium botuli
32	53.4	3.4	1547	1 V30581	Clostridium botuli
33	53.4	3.4	1463	1 V30584	Clostridium botuli
34	53.4	3.4	1472	1 V30585	Clostridium botuli
35	53.4	3.4	1502	1 V30588	Clostridium botuli
36	53.4	3.4	1402	1 V30572	Clostridium botuli
37	46.8	3.0	104	1 T33359	Ich-2 gene 5' prim
38	45	2.9	21591	1 X13047	Enterococcus faec
39	42	2.7	110000	1 X20248_01	Continuation (2 of
40	42	2.7	110000	1 X20248_02	Continuation (3 of
41	40	2.6	110000	1 V21209_15	Continuation (16 o
42	39	2.5	4435	1 Q29686	CCV-CS4 spike gene
43	38.8	2.5	1116	1 O52273	Soybean seed stora

## ALIGNMENTS

RESULT	ID	Query Match	Length	ID	Description
1	T38249	94.0	1530	1 T38249	Y. pestis Fl/V ant
2	T38249	90.0	1462	1 T38256	Y. pestis Fl/V ant
3	T38249	62.4	1014	1 T38222	Partial LcrV (V an
4	T38249	62.3	1014	1 T38242	Y. pestis lcrV (V
5	T38249	62.2	1014	1 T38243	Y. pestis lcrV (V
6	T38249	62.1	1014	1 T04223	Partial LcrV (V an
7	T38249	33.0	544	1 V41594	Nucleotide sequenc
8	T38249	33.0	547	1 O92819	Yersinia pestis ca
9	T38249	33.0	544	1 T38248	Y. pestis Fl antige
10	T38249	33.0	544	1 V41595	Nucleotide sequenc
11	T38249	32.6	474	1 V41600	Nucleotide sequenc
12	T38249	28.9	474	1 V41600	Nucleotide sequenc
13	T38249	28.9	541	1 O92817	Yersinia pestis ca
14	T38249	28.9	541	1 T38244	Y. pestis caFl (Fl
15	T38249	28.8	542	1 O92818	Yersinia pestis ca
16	T38249	28.6	450	1 V41601	Nucleotide sequenc
17	T38249	28.5	447	1 V41609	Nucleotide sequenc
18	T38249	25.7	576	1 V41598	Nucleotide sequenc
19	T38249	21.6	513	1 V41599	Nucleotide sequenc
20	T38249	4.4	861	1 O81500	sfy anti-rev seque
21	T38249	4.4	861	1 T45347	Single chain sfy a
22	T38249	4.4	375	1 V54056	DNA fragment 2 of
23	T38249	4.4	375	1 V54057	DNA fragment 3 of
24	T38249	4.4	375	1 V54058	DNA fragment 4 of
25	T38249	3.4	1546	1 V30575	Clostridium botuli
26	T38249	3.4	1402	1 T29246	Type A neurotoxin
27	T38249	3.4	1278	1 V41263	Chlamydomonas rein
28	T38249	3.4	1469	1 V30591	Clostridium botuli
29	T38249	3.4	1460	1 V30593	Clostridium botuli
30	T38249	3.4	1535	1 V30596	Clostridium botuli
31	T38249	3.4	1526	1 V30580	Clostridium botuli
32	T38249	3.4	1547	1 V30581	Clostridium botuli
33	T38249	3.4	1463	1 V30584	Clostridium botuli
34	T38249	3.4	1472	1 V30585	Clostridium botuli
35	T38249	3.4	1502	1 V30588	Clostridium botuli
36	T38249	3.4	1402	1 V30572	Clostridium botuli
37	T38249	3.0	104	1 T33359	Ich-2 gene 5' prim
38	T38249	2.9	21591	1 X13047	Enterococcus faec
39	T38249	2.7	110000	1 X20248_01	Continuation (2 of
40	T38249	2.7	110000	1 X20248_02	Continuation (3 of
41	T38249	2.6	110000	1 V21209_15	Continuation (16 o
42	T38249	2.5	4435	1 Q29686	CCV-CS4 spike gene
43	T38249	2.5	1116	1 O52273	Soybean seed stora

[illegible]

Accession	Score	Length	Mismatches	Indels	Gaps
D01271	1462	1462	0	0	0
Q01379	1438	1438	0	0	0
D01331	1390	1390	0	0	0
Q01439	1498	1498	0	0	0
D01391	1450	1450	0	0	0
Q01499	1558	1558	0	0	0
D01451	1510	1510	0	0	0
Q01559	1566	1566	0	0	0
D01511	1518	1518	0	0	0
RESULT	2				
ID	T38256	T38256			
AC	T38256	standard; DNA; 1462 BP.			
DT	28-DEC-1996	(first entry)			
DE	Y. pestis F1/V antigen gene fusion.				
KW	Plague; vaccine; genetic immunisation; V antigen; lcrv;				
KW	F1 antigen; calf; ds.				
OS	Chimeric Yersinia pestis strain GB;				
OS	Chimeric synthetic.				
FT	Key	Location/Qualifiers			
FT	cds	8..1450			
FT	/*tag= a	/product= F1/V fusion protein			
FT	452..472	/note= "bases 452-472 is a sequence derived			
FT	misc_feature	from PCR primers"			
FT	/*tag= b				
FT	/*note=				
FT	from PCR primers"				
PN	W09628551.A1.				
PD	19-SEP-1996.				
PR	13-MAR-1996; G00571.				
PR	13-MAR-1995; GB-005059.				
PR	15-SEP-1995; GB-018946.				
PR	05-DEC-1995; GB-024825.				
PR	(MNA ) UK SEC FOR DEFENCE.				
PI	Bennett AM, Leary SEC, Oyston PCF, Titball RW, Williamson ED;				
PI	WPI: 96-433824/43.				
DR	P-PSDB: W01045.				
DR	Yersinia pestis V antigen and F1 antigen or their protective				
PT	epitopic parts - useful in vaccine for protection against plague				
PS	Disclosure; Page 51-55; 98pp; English.				
CC	A gene fusion (T38256) comprises coding sequences for the Yersinia				
CC	pestis F1 antigen (see also T38244) (without the signal sequence)				
CC	and for the Y. pestis V antigen (see also T38243), joined by a				
CC	sequence encoding a 6-amino acid peptide linker. It was obcd. by				
CC	PCR amplification of the individual genes using primers (see also				
CC	T38245) based on the genes and including linker sequences. The				
CC	gene fusion (see also T38249) can be used to produce F1/V fusion				
CC	protein (W01045) in transformed cells, esp. gut-colonising				
CC	organisms, to induce an immune response against Y. pestis, the				
CC	causative organism of plague.				
SQ	Sequence 1462 BP; 476 A; 285 C; 301 G; 400 T;				
Query Match	90.0%; Score 1409; DB 1; Length 1462;				
Best Local Similarity	99.0%; Pred. No. 0;				
Matches 1431; Conservative	0; Mismatches 5; Indels 9; Gaps 1;				
Q01331	CGCGAGATTATACGCAAGCACCACCTGCAAGCGCACTCTGTGTAACGAGCCGCATCA	190			
D01331	CGCGAGATTATACGCAAGCACCACCTGCAAGCGCACTCTGTGTAACGAGCCGCATCA	65			
Q01511	CTTTACATATTAAGGAAGCGCTCCCAATTACATATTATGAGCAATGAAACATCGATACAG	250			

```
Db 66 CTTTACATATAAGSAGGCGCTCCATTACATTTATGCAATGGAACATCGATCAG 125
QY 251 AATTACTTGTGGTAGCGTTACTCTTGGCGGCTATATAACAGGACCACTAGCACATCTG 310
Db 126 AATTACTTGTGGTAGCGTTACTCTTGGCGGCTATATAACAGGACCACTAGCACATCTG 185
QY 311 TTAACCTTACAGATGGCGGGGATGCCATGATCTTAACATTTACTTCTCAGAGATGGA 370
Db 186 TTACACTTACAGATGGCGGGGATGCCATGATCTTAACATTTACTTCTCAGAGATGGA 245
QY 371 ATAACCAACAATTCATACAAAAGATGTTGGCAAGATTTCTAGAGATTTGATATCTTC 430
Db 246 ATACCAACAATTCATACAAAAGATGTTGGCAAGATTTCTAGAGATTTGATATCTTC 305
QY 431 CTAAAGTAAACGGTGAACACCTTGTGGGGGATGACGTGCTTGGCTAGCGGACGACG 490
Db 306 CTAAAGTAAACGGTGAACACCTTGTGGGGGATGACGTGCTTGGCTAGCGGACGACG 365
QY 491 ATTTCTTTGGCTCAATTTGGTCCAAAGCGGTAACCTTGACGACGTAATACACTG 550
Db 366 ATTTCTTTGGCTCAATTTGGTCCAAAGCGGTAACCTTGACGACGTAATACACTG 425
QY 551 ATGCTTAACCCGTAACCGTATCTAACCAAGATTCAT-----GATTAGCGCTAG 601
Db 426 ATGCTTAACCCGTAACCGTATCTAACCAAGATTCATCGAAGTCTGATTAGCGCTAG 485
QY 602 AACCAAAACCCACACATTTTATTGAGGATCTAGAAAAGTTAGGTGAACAACCTACTG 661
Db 486 AACCAAAACCCACACATTTTATTGAGGATCTAGAAAAGTTAGGTGAACAACCTACTG 545
QY 662 GTCATGTTCTTCAAGTTTATGAAAGATTTGGTTCAGTTAGTCAAAAGATATAAGATA 721
Db 546 GTCATGTTCTTCAAGTTTATGAAAGATTTGGTTCAGTTAGTCAAAAGATATAAGATA 605
QY 722 TTTTCATTAATATGATCCCAAGAAAGATTCGGAGGTTTTGGCCATPAGATTAATACG 781
Db 606 TTTTCATTAATATGATCCCAAGAAAGATTCGGAGGTTTTGGCCATPAGATTAATACG 665
QY 782 ATGATATGCAATTTGTCAGAAAGAAATCTAGCTTATTTTCCACCGAGATACCATTTCTA 841
Db 666 ATGATATGCAATTTGTCAGAAAGAAATCTAGCTTATTTTCCACCGAGATACCATTTCTA 725
QY 842 AAGCGGCTATTTATGCAACCAACCTCAAAATGGCATCAAGCAGATGTAAGATTCCTTG 901
Db 726 AAGCGGCTATTTATGCAACCAACCTCAAAATGGCATCAAGCAGATGTAAGATTCCTTG 785
QY 902 AATCATGCCGGAATACACATGGGAATTCGGGCGTTCATGCGAGTAATGATTTCTCTT 961
Db 786 AATCATGCCGGAATACACATGGGAATTCGGGCGTTCATGCGAGTAATGATTTCTCTT 845
QY 962 TAAACCCGATGATGATGATGATGATTTTGAAGATGATTTGATTTCAATGATCATC 1021
Db 846 TAAACCCGATGATGATGATGATGATTTTGAAGATGATTTGATTTCAATGATCATC 905
QY 1022 ATGATGATCCCGGTAGCAAGTTCCGTGAAGAAATAGCTGAGCTTACCGCGGAATTAAG 1081
Db 906 ATGATGATCCCGGTAGCAAGTTCCGTGAAGAAATAGCTGAGCTTACCGCGGAATTAAG 965
QY 1082 TTTATTCAGTTATTCAGCGGAATTAATPAGATCTGCTATAGTATGGCACATTAATA 1141
Db 966 TTTATTCAGTTATTCAGCGGAATTAATPAGATCTGCTATAGTATGGCACATTAATA 1025
QY 1142 TCCATGATTAATTCATTAATCTGATGATTAATAATTTATATGTTTTPACGATGAAGAG 1201
Db 1026 TCCATGATTAATTCATTAATCTGATGATTAATAATTTATATGTTTATACGATGAAGAG 1085
QY 1202 TTTTAAAGCGACGAGAGTACAAAATTTCTCGAGAAAATGCTTCAACCAACCATTCAG 1261
Db 1086 TTTTAAAGCGACGAGAGTACAAAATTTCTCGAGAAAATGCTTCAACCAACCATTCAG 1145
QY 1262 TGGATGAGGAGGAGAAAATATAGTCTCGATTAAGAGACTTTCTTGAAGTGAAGATAAAA 1321
Db 1146 TGGATGAGGAGGAGAAAATATAGTCTCGATTAAGAGACTTTCTTGAAGTGAAGATAAAA 1205
```

```
QY 1322 GAACGGGGCGTTGGGTAATCTGAAAAACTCATACTTTATATAAGATAATTAAGAT 1381
Db 1206 GAACGGGGCGTTGGGTAATCTGAAAAACTCATACTTTATATAAGATAATTAAGAT 1265
QY 1382 TATTCACCTTTCACACACACCTCTCGGATAAATCCAGCGCGCTCAACGACTGGTACG 1441
Db 1266 TATTCACCTTTCACACACACCTCTCGGATAAATCCAGCGCGCTCAACGACTGGTACG 1325
QY 1442 AAAAAACAATCAGCTGCTGATATTACATCAGTTTATTCAGCTATTTGAAGCACTGA 1501
Db 1326 AAAAAACAATCAGCTGCTGATATTACATCAGCTTTTATTCAGCTATTTGAAGCACTGA 1385
QY 1502 ACCGTTTCATCAGAAATATGATTCAGTATGCAACGCTGCTAGATGACAGCTGTGTA 1561
Db 1386 ACCGTTTCATCAGAAATATGATTCAGTATGCAACGCTGCTAGATGACAGCTGTGTA 1445
QY 1562 AATGA 1566
Db 1446 AATGA 1450

RESULT 3
T04222
ID T04222 standard; DNA; 1014 BP.
AC T04222;
DT 18-APR-1996 (first entry)
DE Partial lcrV (V antigen) gene of Y. pestis.
KW lcrV; V antigen; virulence; plague; vaccine; epitope; ss.
OS Yersinia pestis.
FH key
FT cds
FT 1..990
FT /tag= a
FT /note= "V antigen"
PN W09524475-AL.
PD 14-SEP-1995.
PF 06-MAR-1995; G00481.
PR 08-MAR-1994; GB-004577.
PA (MINA) UK SEC FOR DEFENCE.
PI Leary SEC, Tithall RW, Williamson ED, Leary SE;
DR WPT_95-328268/42.
DR P-PSDB; R/9961.
PT Recombinant DNA expressing Yersinia pestis V antigen - useful in
PT oral or parenteral vaccines for protection against plague
PS Claim 6; Page 11-13; 25pp; English.
CC T04222-23 are DNA sequences (lcrV) encoding all or a protective epitopic
CC part of the mature V protein of Yersinia pestis. The protein was
CC expressed as a fusion protein with maltose binding protein or
CC glutathione-S-transferase in 3 different plasmid vectors. Y. pestis is
CC the highly virulent causative organism of plague in a wide range of
CC animals, including man. The V antigen (lcrV) is an unstable 37.3 kDa
CC monomeric peptide encoded on the ca. 70 kb lcr plasmid. The V antigen
CC is postulated to act as a virulence antigen, and transformed
CC microorganisms conty. recombinant DNA encoding a V antigen protein/
CC peptide are useful in vaccines to protect against plague.
SQ Sequence 1014 BP; 346 A; 181 C; 201 G; 286 T;
```

```
Query Match 62.4%; Score 976.8; DB 1: Length 1014;
Best Local Similarity 99.8%; Pred. No. 7.2e-251;
Matches 978; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 587 TGATTAGACCTTACGAAACCAACCAACATTTTATGAGATCTAGAAAAAGTTAGG 646
Db 11 TCATTAGACCTTACGAAACCAACCAACATTTTATGAGATCTAGAAAAAGTTAGG 70
QY 647 TCGAACAACTTCTGCTGCTGCTTCTAGTTTGAAGATTTGTTAGTCAAG 706
Db 71 TCGAACAACTTCTGCTGCTGCTTCTAGTTTGAAGATTTGTTAGTCAAG 130
QY 707 ATPAAAAATATATATTTCCATTAATATATGATCCAGAAAAATTCGAGATTTTGCCA 766
Db 131 ATPAAAAATATATATTTCCATTAATATATGATCCAGAAAAATTCGAGATTTTGCCA 190
```

```
QY 767 ATAGAGTAATTGATGATATGGAATGCTCAAGAAAATCCTAGCTATTATTTCTACCG 826
DB 191 ATAGAGTAATTGATGATATGGAATGCTCAAGAAAATCCTAGCTATTATTTCTACCG 250
QY 827 AGATACCAATCTTAAAGGGCGTCATATGACAACCAACTGCAAAAATGGCATCAAGCGAG 886
DB 251 AGGATGCGATCTTAAAGGGCGTCATATGACAACCAACTGCAAAAATGGCATCAAGCGAG 310
QY 887 TAAAGAGTTCCTGAATCATCGCCGAATACACAAATGGGAATGGCGGCTTCATGGCAG 946
DB 311 TAAAGAGTTCCTGAATCATCGCCGAATACACAAATGGGAATGGCGGCTTCATGGCAG 370
QY 947 TAATGCAATTCCTTTTACCGCGCATGCTATCATGATGATATTTTAAAGTGAATGCTTG 1006
DB 371 TAATGCAATTCCTTTTACCGCGCATGCTATCATGATGATATTTTAAAGTGAATGCTTG 430
QY 1007 ATTCATGATCATCATGATGATGCGCGTACGAAGTGGCTGAAGAAATTAAGCTGAGCTTA 1066
DB 431 ATTCATGATCATCATGATGATGCGCGTACGAAGTGGCTGAAGAAATTAAGCTGAGCTTA 490
QY 1067 CCGCCGAATTAAGATTTATTCAGTTATTCAGCCGAATTAATTAAGCATCTGCTAGTA 1126
DB 491 CCGCCGAATTAAGATTTATTCAGTTATTCAGCCGAATTAATTAAGCATCTGCTAGTA 550
QY 1127 GTGGCACCAATTAATTCATGATATTCATTCATTCATGATGATGATGATGATGATGAT 1186
DB 551 GTGGCACCAATTAATTCATGATATTCATTCATTCATTCATGATGATGATGATGATGATG 610
QY 1187 ATACAGATGAGAGATTTTAAAGCCAGCCAGATCAAAATTCCTGAGAAATGCTC 1246
DB 611 ATACAGATGAGAGATTTTAAAGCCAGCCAGATCAAAATTCCTGAGAAATGCTC 670
QY 1247 AAACCAACCATTCAGTGATGAGGAGCGAGAAAAAATAGTCTGATTAAGAGACTTCTTG 1306
DB 671 AAACCAACCATTCAGTGATGAGGAGCGAGAAAAAATAGTCTGATTAAGAGACTTCTTG 730
QY 1307 GAAGTGAATTAAGAACCGGGCGCTGGGTATCTGAAAAAATCTACTCTTATATA 1366
DB 731 GAAGTGAATTAAGAACCGGGCGCTGGGTATCTGAAAAAATCTACTCTTATATA 790
QY 1367 AAGATATATGATATTCCTACTTCCACCACTGCTGGATTAAGTCCAGGCGCTCA 1426
DB 791 AAGATATATGATATTCCTACTTCCACCACTGCTGGATTAAGTCCAGGCGCTCA 850
QY 1427 AGCACTTGTTAGCCAAAAACAACACTGCTGCTGATTAATCACTGTTTAAATTCAG 1486
DB 851 AGCACTTGTTAGCCAAAAACAACACTGCTGCTGATTAATCACTGTTTAAATTCAG 910
QY 1487 CTATGAGACACTGACCGCTTTCATTGAGAAATATGATTCAGATGATCAACGCTGCTAG 1546
DB 911 CTATGAGACACTGACCGCTTTCATTGAGAAATATGATTCAGATGATCAACGCTGCTAG 970
QY 1547 ATGACAGCTGCTGTAATGA 1566
DB 971 ATGACAGCTGCTGTAATGA 990

RESULT 4
ID T38242
AC T38242;
DT 28-DEC-1996 (first entry)
DE Y. pestis lcrV (V antigen) gene.
KW Plaque; vaccine; genetic immunisation; V antigen; lcrV;
OS Yersinia pestis strain GB.
FH key location/Qualifiers
FT cds 1..990
FT 1..tag- a
FT 1..misc- feature 1..6
FT 1..tag- b
FT /note- "vector pMAL-p2 or pMAL-c2-derived bases"
```

```
FT mutation 1006
FT /tag= C
FT /note- "base 1006 is altered to a T to create a
DB second in-frame stop codon"
PN W09628551-1.
PD 19-SEP-1996.
PR 13-MAR-1996; GB-0050571.
PR 13-MAR-1995; GB-005059.
PR 15-SEP-1995; GB-018946.
PR 05-DEC-1995; GB-024825.
PA (MINA ) UK SEC FOR DEFENCE.
PI Bennett AM, Leary SEC, Oyston PCF, Tibball RW, Williamson ED;
DR MPI: 96-433824/43.
DR P-PSDB: W01040.
PT Yersinia pestis V antigen and F1 antigen or their protective
PT epitopic parts - useful in vaccine for protection against plague
PS Claim 5; Page 25-28; 98bp; English.
CC A lcrV gene sequence (T38242) codes for the Yersinia pestis V
CC antigen (W01040), which is capable of evoking protective immune
CC responses in animals. The gene was amplified from Y. pestis
CC DNA by PCR using primers (T38250-51) homologous to the 5' and 3'
CC ends of the gene. The gene was inserted into vector pMAL-p2,
CC pMAL-c2 or pGEX-5X-2 (see also T38243) to allow prodn. of
CC recombinant V antigen for use in vaccines against plague.
CC Expression in gut-colonising organisms and attenuated Salmonella
CC typhn allows prodn. of live vaccines. F1/V antigen fusions were
CC also created (see also T38249 and T38256). The gene can itself be
CC used in genetic vaccines.
SQ Sequence 1014 BP; 347 A; 180 C; 201 G; 286 T;
```

Query Match 62.3%; Score 975.2; DB 1; Length 1014;  
Best Local Similarity 99.7%; Pred. No. 1.9e-250;  
Matches 977; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 587 TGAATTGAGCGCTGCGAACAACCCCAACATTTTGTGAGATCTAGAAAAAGTAGGG 646
DB 11 TCAATTGAGCGCTGCGAACAACCCCAACATTTTGTGAGATCTAGAAAAAGTAGGG 70
QY 647 TGGACAACTTACTGCTGATGTTGTTCTTCAAGTTTAAAGAAATGGTTCAGTTAGTCAAG 706
DB 71 TGGACAACTTACTGCTGATGTTGTTCTTCAAGTTTAAAGAAATGGTTCAGTTAGTCAAG 130
QY 707 ATAAAAATATGATATTTCCATTAATATGATGCCAGAAAAGATTCGAGGTTTTGCCA 766
DB 131 ATAAAAATATGATATTTCCATTAATATGATGCCAGAAAAGATTCGAGGTTTTGCCA 190
QY 767 ATGAGTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 826
DB 191 ATGAGTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 250
QY 827 AGGATACCAATCTTAAAGGGCGTCATATGACAACCAACTGCAAAAATGGCATCAAGCGAG 886
DB 251 AGGATGCGATCTTAAAGGGCGTCATATGACAACCAACTGCAAAAATGGCATCAAGCGAG 310
QY 887 TAAAGAGTTCCTGAATCATCGCCGAATACACAAATGGGAATGGCGGCTTCATGGCAG 946
DB 311 TAAAGAGTTCCTGAATCATCGCCGAATACACAAATGGGAATGGCGGCTTCATGGCAG 370
QY 947 TAATGCAATTCCTTTTACCGCGCATGCTATCATGATGATATTTTAAAGTGAATGCTTG 1006
DB 371 TAATGCAATTCCTTTTACCGCGCATGCTATCATGATGATATTTTAAAGTGAATGCTTG 430
QY 1007 ATTCATGATCATCATGATGATGCGCGTACGAAGTGGCTGAAGAAATTAAGCTGAGCTTA 1066
DB 431 ATTCATGATCATCATGATGATGCGCGTACGAAGTGGCTGAAGAAATTAAGCTGAGCTTA 490
QY 1067 CCGCCGAATTAAGATTTATTCAGTTATTCAGCCGAATTAATTAAGCATCTGCTAGTA 1126
DB 491 CCGCCGAATTAAGATTTATTCAGTTATTCAGCCGAATTAATTAAGCATCTGCTAGTA 550
QY 1127 GTGGCACCAATTAATTCATGATATTCATTCATTCATTCATGATGATGATGATGATGATG 1186
```



DB 551 GTGCACCATTAATATCCATGATTAATTCATTAATCTCTGATATAAAATTTATATGTT 610  
QY 1187 ATACAGATGAAGAGATTTTAAAGCCAGCCAGATACAAAATTCGAGAAATGCCCTC 1246  
DB 611 ATACAGATGAAGAGATTTTAAAGCCAGCCAGATACAAAATTCGAGAAATGCCCTC 670  
QY 1247 AAACCAACCATTCAGTGGTGGAGCGAGAAAAAATAGTCTCGATTAAGACCTTCTTG 1306  
DB 671 AAACCAACCATTCAGTGGTGGAGCGAGAAAAAATAGTCTCGATTAAGACCTTCTTG 730  
QY 1307 GAAGGAGATTAAGAACCGGGGGTGGTATCTGAAAACATCATCTCTTAATAA 1366  
DB 731 GAAGGAGATTAAGAACCGGGGGTGGTATCTGAAAACATCATCTCTTAATAA 790  
QY 1367 AAGATAATATGATTAATCTACTTTGCGACCACTGCTCGGATTAAGTCCAGCGCTCA 1426  
DB 791 AAGATAATATGATTAATCTACTTTGCGACCACTGCTCGGATTAAGTCCAGCGCTCA 850  
QY 1427 ACGACTTGGTTAGCCAAAACAACTCAGCTGCTGATATTAATCAGCTTTAAATTCAG 1486  
DB 851 ACGACTTGGTTAGCCAAAACAACTCAGCTGCTGATATTAATCAGCTTTAAATTCAG 910  
QY 1487 CTAATGAACCATGACCGCTTCATTCAGAAATATGATTCAGTATGACAGCTGCTAG 1546  
DB 911 CTAATGAACCATGACCGCTTCATTCAGAAATATGATTCAGTATGACAGCTGCTAG 970  
QY 1547 ATGACAGCTGCTGTAATGA 1566  
DB 971 ATGACAGCTGCTGTAATGA 990

RESULT 5  
T38243  
ID T38243 standard; DNA: 1014 BP.  
AC T38243: 1996 (first entry)  
DE 28-DEC-1996 (first entry)  
DE Y. pestis lcrv (V antigen) gene.  
KW Plague; vaccine; genetic immunisation; V antigen; lcrv;  
KW FI antigen; ds.  
OS Yersinia pestis strain GB.  
FH Key location/Qualifiers  
FT cds 1..990  
FT misc\_feature 1..10  
FT /tag= a  
FT /note= "vector pGEX-5X-2-derived bases"  
FT mutation 16  
FT /tag= c  
FT /note= "base 16 is altered to a C from an A to create an EcoRI site"  
FT mutation 1006  
FT /tag= d  
FT /note= "base 1006 is altered to a T to create a second in-frame stop codon"  
PN W09628551-A1.  
PD 19-SEP-1996.  
PE 13-MAR-1996: G00571.  
PR 13-MAR-1995: GB-005059.  
PR 15-SEP-1995: GB-018946.  
PR 05-DEC-1995: GB-024825.  
PA (MIMA) UK SEC FOR DEFENCE.  
PI Bennett AM, Leary SEC, Oyston PCF, Tilball RM, Williamson ED;  
DR WPI: 96-433824/43.  
P-BSDB: W01041.  
PT Yersinia pestis V antigen and FI antigen or their protective  
PT epitopic parts - useful in vaccine for protection against plague  
PS Claim 5; Page 32-35; 98BP: English.  
CC A lcrv gene sequence (T38243) codes for the Yersinia pestis V  
CC antigen (W01041), which is capable of evoking protective immune  
CC responses in animals. The gene was amplified from Y. pestis  
CC DNA by PCR using primers (T38251 and T38259) homologous to the 5'  
CC and 3' ends of the gene. The gene was inserted into vector  
CC pGEX-5x-2, pMAL-p2 or pMAL-c2 (see also T38242) to allow prodn. of

CC recombinant V antigen for use in vaccines against plague.  
CC Expression in gut-colonising organisms and attenuated Salmonella  
CC typhi allows live vaccine prodn. FI/V antigen fusions were also  
CC created (see also T38249 and T38256). The gene can itself be used  
CC in genetic vaccines.  
SQ Sequence 1014 BP; 343 A; 184 C; 205 G; 282 T;  
Query Match 62.2%; Score 974.8; DB 1; Length 1014;  
Best Local Similarity 99.8%; Pred. No. 2.5e-250;  
Matches 976; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 589 ATTAGAGCTTACGAAACAAACCCACACATTTTATGAGATCTGAAAAAGTTAGGGTG 648  
DB 13 ATTGAGGCTACGAAACAAACCCACACATTTTATGAGATCTGAAAAAGTTAGGGTG 72  
QY 649 GAACCACTTACTGTCATGCTGTTCTTCACTTTTGAAGAATTGTTCACTAGTCAAGAT 708  
DB 73 GAAACAATCTACTGTCATGCTGTTCTTCACTTTTGAAGAATTGTTCACTAGTCAAGAT 132  
QY 709 AAAAATATGATATTTCCATTAATATATGATCCAGAAAAAGATTGGAGTTTTCCTAAT 768  
DB 133 AAAAATATGATATTTCCATTAATATATGATCCAGAAAAAGATTGGAGTTTTCCTAAT 192  
QY 769 AGAGTAATTAATGATATGATATGATATGATATGATATGATATGATATGATATGATAT 828  
DB 133 AGAGTAATTAATGATATGATATGATATGATATGATATGATATGATATGATATGATAT 252  
QY 829 GATACCAATTTTAAAGGGGCTATTATGACAAACCACTGCAAAATGCGATCAAGCGAGTA 888  
DB 253 GATGCCATTTTAAAGGGGCTATTATGACAAACCACTGCAAAATGCGATCAAGCGAGTA 312  
QY 889 AAAGAGTTCCTTGATATGATATGATATGATATGATATGATATGATATGATATGATAT 948  
DB 313 AAAGAGTTCCTTGATATGATATGATATGATATGATATGATATGATATGATATGATAT 372  
QY 949 ATGCAATTTCTCTTAAAGCGCGCATGCTATGATATGATATGATATGATATGATATGAT 1008  
DB 373 ATGCAATTTCTCTTAAAGCGCGCATGCTATGATATGATATGATATGATATGATATGAT 432  
QY 1009 TCAATGAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1068  
DB 433 TCAATGAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 492  
QY 1069 GCGGAATTAAGATTTTATGATATGATATGATATGATATGATATGATATGATATGAT 1128  
DB 493 GCGGAATTAAGATTTTATGATATGATATGATATGATATGATATGATATGATATGAT 552  
QY 1129 GGCACCATTAATATGATGATATGATATGATATGATATGATATGATATGATATGAT 1188  
DB 553 GGCACCATTAATATGATGATATGATATGATATGATATGATATGATATGATATGATAT 612  
QY 1189 ACAGATGAAGAGATTTTAAAGCCAGCCAGATACAAAATTCGAGAAATGCCCTCA 1248  
DB 613 ACAGATGAAGAGATTTTAAAGCCAGCCAGATACAAAATTCGAGAAATGCCCTCA 672  
QY 1249 ACCACCATTCAGTGGTGGAGCGAGAAAAAATAGTCTGATTAAGAGACTTTTGGGA 1308  
DB 673 ACCACCATTCAGTGGTGGAGCGAGAAAAAATAGTCTGATTAAGAGACTTTTGGGA 732  
QY 1309 AGTGAATTAAGAAACCGGGGGTGGTATCTGAAAACATCATCTCTTAATAA 1368  
DB 733 AGTGAATTAAGAAACCGGGGGTGGTATCTGAAAACATCATCTCTTAATAA 792  
QY 1369 GATAATTAATGATTTATCTACTTTGCGACCACTGCTCGATTAAGTCCAGCGCTCAAC 1428  
DB 793 GATAATTAATGATTTATCTACTTTGCGACCACTGCTCGATTAAGTCCAGCGCTCAAC 852  
QY 1429 GACTTGGTTAGCCAAAACAACTCAGCTGCTGATATTAATCAGCTTTTAATTCAGCT 1488  
DB 853 GACTTGGTTAGCCAAAACAACTCAGCTGCTGATATTAATCAGCTTTTAATTCAGCT 912  
QY 1489 ATTGAAGCACTGAACCGTTTCAATTAGAAATATGATTCAGTATGCAACGCTGCTAAT 1548

```
|||||
Db 913 ATTGAAGCAGCTGAACCGTTTCATTACAGAAATATGATTCACTGATGACAGCTCTGATGAT 972
OY 1549 GACACGCTCTGTGTAATGA 1566
   |||||||
Db 973 GACACGCTCTGTGTAATGA 990

RESULT 6
T04223
ID T04223 standard; DNA; 1014 BP.
AC T04223:
DE 18-APR-1996 (first entry)
DE Partial LcrV (V antigen) gene of Y. pestis.
KW LcrV; V antigen; virulence; plague; vaccine; epitope; ss.
OS Yersinia pestis.
FH Key Location/Qualifiers
FT cds 1..990
   /tag= a
PN WO9524475-A1.
PD 14-SEP-1995.
PE 06-MAR-1995; G00481.
PR 08-MAR-1994; GB-004577.
PA (MINA ) UK SEC FOR DEFENCE.
PI Leary SEC, Tibball RW, Williamson ED, Leary SE;
DR WPI: 95-328268/42.
P-PSDB: R79962.
PT Recombinant DNA expressing Yersinia pestis V antigen - useful in
PT oral or parenteral vaccines for protection against plague
PS Claim 6: Page 15-16: 25pp: English.
CC T04222-23 are DNA sequences (LcrV) encoding all or a protective epitopic
CC part of the mature V protein of Yersinia pestis. The protein was
CC expressed as a fusion protein with maltose binding protein or
CC glutathione-S-transferase in 3 different plasmid vectors. Y. pestis
CC is the highly virulent causative organism of plague in a wide range of
CC animals, including man. The V antigen (LcrV) is an unstable 37.3 kDa
CC monomeric peptide encoded on the ca. 70 kb Lcr plasmid. The V antigen
CC is postulated to act as a virulence antigen, and transformed
CC microorganisms contg. recombinant DNA encoding a V antigen protein/
CC peptide are useful in vaccines to protect against plague.
SQ Sequence 1014 BP; 343 A; 185 C; 204 G; 282 T;
```

```
Query Match 62.1%; Score 973.2; DB 1; Length 1014;
Best Local Similarity 99.7%; Pred. No. 6.6e-250;
Matches 975; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 589 ATTAGAGCCTAGCAACAAACCCACACATTTATGAGATCTAGAAAAGTTAGAGTG 648
   |||||||
Db 13 ATTGAGCCTAGCAACAAACCCACACATTTATGAGATCTAGAAAAGTTAGAGTG 72
OY 649 GAACAACCTACTGTCATGGTCTTCAGTTTGAAGAATGGTTCAGTTAGTCAAAAGAT 708
   |||||||
Db 73 GAACAACCTACTGTCATGGTCTTCAGTTTGAAGAATGGTTCAGTTAGTCAAAAGAT 132
OY 709 AAAAATATAGATATTTCCATTAAATATGATCCAGAAAAGATTCGAGGTTTGGCCAAT 768
   |||||||
Db 133 AAAAATATAGATATTTCCATTAAATATGATCCAGAAAAGATTCGAGGTTTGGCCAAT 192
OY 769 AGAGTAATTACTGATATATCGAATGCTCAAGAAAATCCTACTATTATTTCACCCGAG 828
   |||||||
Db 193 AGAGTAATTACTGATATATCGAATGCTCAAGAAAATCCTACTATTATTTCACCCGAG 252
OY 829 GATACCATTTCTTAAAGCGGCTTATGACAACCACTGCAAAATGGATCAAGCAGATGA 888
   |||||||
Db 253 GATGCCATTTCTTAAAGCGGCTTATGACAACCACTGCAAAATGGATCAAGCAGATGA 312
OY 889 AAAGACTTCTTGAATCATGCGGAATACCAATGGGAATGGCGGCTTCATGGCAGTA 948
   |||||||
Db 313 AAAGACTTCTTGAATCATGCGGAATACCAATGGGAATGGCGGCTTCATGGCAGTA 372
OY 949 ATGCATTTCTCTTAACCGCGGATCGATGATGATATTTTGAAGAAGATGTTGAT 1008
   |||||||

|||||
Db 373 ATGCATTTCTCTTAAACCCCGGATCGATGATGATATTTTGAAGATGATGTTGAT 432
OY 1009 TCAATGAATCATCATGATGATGATGCCGTAGCAAGTTGCGTGAAGAATTAGCAGCTTACC 1068
   |||||||
Db 433 TCAATGAATCATCATGATGATGATGCCGTAGCAAGTTGCGTGAAGAATTAGCAGCTTACC 492
OY 1069 GCCGAATTAAAGATTTTATTCAGTTATTCAGCCGGAATTAATAGCATCTGCTAGTACT 1128
   |||||||
Db 493 GCCGAATTAAAGATTTTATTCAGTTATTCAGCCGGAATTAATAGCATCTGCTAGTACT 552
OY 1129 GGCACCATTAATTCATGATTAATCCATTAAATTCATGATGATTAATTAATGATGTTAT 1188
   |||||||
Db 553 GGCACCATTAATTCATGATTAATCCATTAAATTCATGATGATTAATTAATGATGTTAT 612
OY 1189 ACAGATGAAGAGATTTTAAAGCAGCAGCAGAGTACAAATTTCTCGAGAAATGCGCTCAA 1248
   |||||||
Db 613 ACAGATGAAGAGATTTTAAAGCAGCAGCAGAGTACAAATTTCTCGAGAAATGCGCTCAA 672
OY 1249 ACCACCATTCAGTGTGATGGAGACGAGAAAATAATGTTCTGATTAAGGACTTTCTGGA 1308
   |||||||
Db 673 ACCACCATTCAGTGTGATGGAGACGAGAAAATAATGTTCTGATTAAGGACTTTCTGGA 732
OY 1309 AGTGAATTAATTAAGAACCGGGGGGTGGTAAATCTGAATAAAGTCACTATATATAA 1368
   |||||||
Db 733 AGTGAATTAATTAAGAACCGGGGGGTGGTAAATCTGAATAAAGTCACTATATATAA 792
OY 1369 GATAATTAATGAATTAATCTGACTTTGCCACACCTGCTCGATAGTCCAGCGCGTCAAC 1428
   |||||||
Db 793 GATAATTAATGAATTAATCTGACTTTGCCACACCTGCTCGATAGTCCAGCGCGTCAAC 852
OY 1429 GACTTGTTAGCCAAACAACTACGCTGCTGATATTATCATCAGCTTTTAATTCAGCT 1488
   |||||||
Db 853 GACTTGTTAGCCAAACAACTACGCTGCTGATATTATCATCAGCTTTTAATTCAGCT 912
OY 1489 ATTGAAGCAGCTGACCGTTCACTGCAATATGATTCGATGCAACGCTGCTAGAT 1548
   |||||||
Db 913 ATTGAAGCAGCTGACCGTTCACTGCAATATGATTCGATGCAACGCTGCTAGAT 972
OY 1549 GACACGCTCTGTGTAATGA 1566
   |||||||
Db 973 GACACGCTCTGTGTAATGA 990

RESULT 7
V41594
ID V41594 standard; DNA; 544 BP.
AC V41594:
DE 26-OCT-1998 (first entry)
DE Nucleotide sequence of FI antigen nYFI(a)sec544.
KW FI antigen; plasmid; vaccine; plague; ds.
OS Yersinia pestis.
FH Key Location/Qualifiers
FT cds 17..529
   /tag= a
   /product= "FI antigen"
PN WO9824912-A2.
PD 11-JUN-1998.
PE 04-DEC-1997; U22617.
PR 04-DEC-1996; US-767115.
PA (HESK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR WPI: 98-333331/29.
P-PSDB: W59782.
PT Protection of animals against plague - using nucleic acid encoding
PT antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 8: Page 51; 75pp; English.
CC This is the nucleotide sequence of a Yersinia pestis FI antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
SQ Sequence 544 BP; 164 A; 120 C; 114 G; 146 T;
```

Query Match 33.0%; Score 517.2; DB 1; Length 544;  
 Best Local Similarity 98.5%; Pred. No. 1,4e-128;  
 Matches 522; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 68 ATATGAAAAAATCAGTCCGTTATGCCATTCGATTATTTGGACATTCGAACTGCTTA 127  
 |||||||  
 DB 15 ATATGAAAAAATCAGTCCGTTATGCCATTCGATTATTTGGACATTCGAACTGCTTA 74

QY 128 ATGCGGAGATTAACTGCAAGCACCCTGCAAGCGCACTCTGTGAACCCGCGCA 187  
 |||||||  
 DB 75 ATGCGGAGATTAACTGCAAGCACCCTGCAAGCGCACTCTGTGAACCCGCGCA 134

QY 188 TCACCTTACATATAAGAGGCGCTCCAAATTAATATGACAAATGAAACATCGATA 247  
 |||||||  
 DB 135 TCACCTTACATATAAGAGGCGCTCCAAATTAATATGACAAATGAAACATCGATA 194

QY 248 CAGAAATTAATGTTGGTACCTTACTCTGGCGGCTATAAAGACGAACTAGACAT 307  
 |||||||  
 DB 195 CAGAAATTAATGTTGGTACCTTACTCTGGCGGCTATAAAGACGAACTAGACAT 254

QY 308 CTGTACTTACATAGAGCGCGGCGGTATGCCATTAATTAATTAATTTCTAGAGATG 367  
 |||||||  
 DB 255 CTGTACTTACATAGAGCGCGGCGGTATGCCATTAATTAATTAATTTCTAGAGATG 314

QY 368 GAATAACACCAATTCATCAACAAAGTATGSCAAGATTCAGAGATTTGATATCT 427  
 |||||||  
 DB 315 GAATAACACCAATTCATCAACAAAGTATGSCAAGATTCAGAGATTTGATATCT 374

QY 428 CTCCTAAGTAAACGGTGAACCTTGTGGGGATGACGTCGTCCTGCTACGCGGACGC 487  
 |||||||  
 DB 375 CTCCTAAGTAAACGGTGAACCTTGTGGGGATGACGTCGTCCTGCTACGCGGACGC 434

QY 488 AGAATTTCTTGTTCGCTCAATGTCCTCAAGGCGGTAACTGACGAGTAAATACA 547  
 |||||||  
 DB 435 AGAATTTCTTGTTCGCTCAATGTCCTCAAGGCGGTAACTGACGAGTAAATACA 494

QY 548 CTGATGCTGTAAACCGTAAACGCTATCTAACCAAGATTCATGATTAGGCC 597  
 |||||||  
 DB 495 CTGATGCTGTAAACCGTAAACGCTATCTAACCAAGATTCATGATTAGGCC 544

RESULT 8  
 Q92819  
 ID 092819 standard: DNA; 542 BP.  
 AC 092819:  
 DT 17-DEC-1995 (first entry)  
 DE Yersinia pestis cafi (F1) antigen in plasmid pFOR1b.  
 KW Vaccine; antigen; Salmonella typhimurium; Salmonella typhi;  
 KM bubonic plague; pneumonic plague; ds.  
 OS Yersinia pestis.  
 FH Key Location/Qualifiers  
 FT cds 2..7  
 FT /\*tag= a  
 FT /note= "first protein encoded by pFOR1b"  
 FT misc\_feature 1..6  
 FT /\*tag= b  
 FT misc\_feature 336..541  
 FT /\*tag= c  
 FT /\*note= "cafi open reading frame downstream seq."  
 FT cds 21..530  
 FT /\*tag= d  
 FT /\*note= "cafi fusion protein"  
 PN W09518231-A1.  
 PD 06-JUL-1995.  
 PF 23-DEC-1994; G02818.  
 PR 24-DEC-1993; GB-026425.  
 PA (MINA) UK SEC FOR DEFENCE.  
 PI Howells A, Leary SEC, Oyston PCF, Titball RW, Williamson ED;  
 DR WPI: 95-246396/32.  
 DR P-PSDB: R76528.  
 PT DNA constructs capable of transforming microorganisms - which can be  
 PT used as live or attenuated vaccines which induce an immune response,

PT against Yersinia pestis, at mucosal surfaces.  
 PS Disclosure: Page 19-20; 27pp: English.  
 CC The sequence represents the plasmid pFOR1b including the entire  
 CC Y. pestis cafi (F1) antigen gene having a 5' tail including a SacI  
 CC restriction site, and up to 1418G downstream of the cafi ORF.  
 CC The DNA construct can be used to transform human or animal gut  
 CC colonizing microorganisms, specifically attenuated Salmonella  
 CC typhimurium or Salmonella typhi. The transformed microorganisms  
 CC can be used as live/attenuated vaccines which induce immune  
 CC responses at mucosal surfaces. The vaccines provide protection  
 CC against infection with Y. pestis, and are parenterally and orally  
 CC active vaccines offering protection against bubonic and pneumonic  
 CC plague.  
 SQ Sequence 542 BP; 164 A; 118 C; 114 G; 146 T;

Query Match 33.0%; Score 516.8; DB 1; Length 542;  
 Best Local Similarity 99.6%; Pred. No. 1,8e-128;  
 Matches 518; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 68 ATATGAAAAAATCAGTCCGTTATGCCATTCGATTATTTGGACATTCGAACTGCTTA 127  
 |||||||  
 DB 19 ATATGAAAAAATCAGTCCGTTATGCCATTCGATTATTTGGACATTCGAACTGCTTA 78

QY 128 ATGCGGAGATTAACTGCAAGCACCCTGCAAGCGCACTCTGTGAACCCGCGCA 187  
 |||||||  
 DB 79 ATGCGGAGATTAACTGCAAGCACCCTGCAAGCGCACTCTGTGTGAACCCGCGCA 138

QY 188 TCACCTTACATATAAGAGGCGCTCCAAATTAATATGACAAATGAAACATCGATA 247  
 |||||||  
 DB 139 TCACCTTACATATAAGAGGCGCTCCAAATTAATATGACAAATGAAACATCGATA 198

QY 248 CAGAAATTAATGTTGGTACCTTACTCTGGCGGCTATAAAGACGAACTAGACAT 307  
 |||||||  
 DB 199 CAGAAATTAATGTTGGTACCTTACTCTGGCGGCTATAAAGACGAACTAGACAT 258

QY 308 CTGTACTTACATAGAGCGCGGCGGTATGCCATTAATTAATTAATTTCTAGAGATG 367  
 |||||||  
 DB 259 CTGTACTTACATAGAGCGCGGCGGTATGCCATTAATTAATTAATTTCTAGAGATG 318

QY 368 GAATAACACCAATTCATCAACAAAGTATGSCAAGATTCAGAGATTTGATATCT 427  
 |||||||  
 DB 319 GAATAACACCAATTCATCAACAAAGTATGSCAAGATTCAGAGATTTGATATCT 378

QY 428 CTCCTAAGTAAACGGTGAACCTTGTGGGGATGACGTCGTCCTGCTACGCGGACGC 487  
 |||||||  
 DB 379 CTCCTAAGTAAACGGTGAACCTTGTGGGGATGACGTCGTCCTGCTACGCGGACGC 438

QY 488 AGAATTTCTTGTTCGCTCAATGTCCTCAAGGCGGTAACTGACGAGTAAATACA 547  
 |||||||  
 DB 439 AGAATTTCTTGTTCGCTCAATGTCCTCAAGGCGGTAACTGACGAGTAAATACA 498

QY 548 CTGATGCTGTAAACCGTAAACGCTATCTAACCAAGATTCATGATTAGGCC 587  
 |||||||  
 DB 499 CTGATGCTGTAAACCGTAAACGCTATCTAACCAAGATTCATGATTAGGCC 538

RESULT 9  
 T38248  
 ID T38248 standard: DNA; 547 BP.  
 AC T38248:  
 DT 28-DEC-1996 (first entry)  
 DE Y. pestis F1 antigen cafi gene (including signal sequence).  
 KW Plague; vaccine; genetic immunisation; F1 antigen; cafi;  
 KM Y antigen; ds.  
 OS Yersinia pestis strain GB.  
 FH Key Location/Qualifiers  
 FT signal\_peptide 29..91  
 FT /\*tag= a  
 FT mat\_peptide 92..538  
 FT /\*tag= b  
 PN W09628551-A1.  
 PD 19-SEP-1996.

PF 13-MAR-1996: G00571.  
PR 13-MAR-1995: GB-005059.  
PR 13-SEP-1995: GB-018946.  
PR 05-DEC-1995: GB-024825.  
PA (MINA ) UK SEC FOR DEFENCE.  
PI Bennett AM, Leary SEC, Oyston PCF, Tibball RW, Williamson ED;  
DR WPI: 96-433824/43.  
DR P-PSDB: W01043.  
PT Yersinia pestis V antigen and F1 antigen or their protective  
epitopic parts - useful in vaccine for protection against plague  
PS Claim 41, Page 61-62; 98pp; English.  
CC A DNA sequence (T38248) comprises the cafI gene, including the  
CC signal sequence, coding for the F1 antigen (W01043) of Yersinia  
CC pestis. It was obd. by PCR amplification (see also T38257-58)  
CC of Y. pestis DNA. The PCR product was cloned into plasmid pBKCMV  
CC and the resulting plasmid (pF1AB) was used to transform E. coli Nova  
CC Blue. Purified plasmid, when administered by i.m. injection,  
CC induced an immunoglobulin response to F1 in BALB/c mice. Live  
CC vaccines comprising gut colonizing organisms transformed with the  
CC cafI gene (see also T38244) can be used to protect a host animal  
CC against plague.  
SQ Sequence 547 BP; 165 A; 120 C; 115 G; 147 T;

Query Match 33.0%; Score 516.8; DB 1; Length 547;  
Best Local Similarity 99.6%; Pred. No. 1.8e-128;  
Matches 518; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 68 ATATGAAAAAATCAGTCCGCTTATCGCCATTGATTTGTAACCTTATGCACTGCTA 127  
DB 27 ATATGAAAAAATCAGTCCGCTTATCGCCATTGATTTGTAACCTTATGCACTGCTA 86  
QY 128 ATGCGGAGATTTAATGCAAGCACACCTGCAAGCGCACTCTGTTGAACAGCCGCA 187  
DB 87 ATGCGGAGATTTAATGCAAGCACACCTGCAAGCGCACTCTGTTGAACAGCCGCA 146  
QY 188 TCACCTTACATATATAGGAAGGCGCTCCATTTACATTTATGCAATATGAAACATCGATA 247  
DB 147 TCACCTTACATATATAGGAAGGCGCTCCATTTACATTTATGCAATATGAAACATCGATA 206  
QY 248 CAGATTACTGTTGGTAGCGCTTACTCTTGCGGCTATTAACAAGAACCACTAGACAT 307  
DB 207 CAGATTACTGTTGGTAGCGCTTACTCTTGCGGCTATTAACAAGAACCACTAGACAT 266  
QY 308 CTGTTAACTTTACAGATGCCGCGGGTGATCCCATGTACTTAACATTTACTTCAGATG 367  
DB 267 CTGTTAACTTTACAGATGCCGCGGGTGATCCCATGTACTTAACATTTACTTCAGATG 326  
QY 368 GAAATTAACCAACCAATTTACTACAAAGATGATGGCAAGGATTTAGAGATTTTGTATCT 427  
DB 327 GAAATTAACCAACCAATTTACTACAAAGATGATGGCAAGGATTTAGAGATTTTGTATCT 386  
QY 428 CTCCTAAGTAAACGGTAGAACCTTGAGGGGATGACGTGCTTGCTAGGGGCAACC 487  
DB 387 CTCCTAAGTAAACGGTAGAACCTTGAGGGGATGACGTGCTTGCTAGGGGCAACC 446  
QY 488 AGGATTTCTTTGTCGCTCAATTTGTTCCAAAGCGGTAAACTTGACAGCAGTAAATACA 547  
DB 447 AGGATTTCTTTGTCGCTCAATTTGTTCCAAAGCGGTAAACTTGACAGCAGTAAATACA 506  
QY 548 CTGATGCTGTAAACCGTATCTATCTAACAAGATTTCAT 587  
DB 507 CTGATGCTGTAAACCGTATCTATCTAACAAGATTTCAT 546

RESULT 10  
V41596  
ID V41596 standard; DNA; 544 BP.  
AC V41596;  
DT 26-OCT-1998 (first entry)  
DE Nucleotide sequence of F1 antigen nypf1sec510.  
KW F1 antigen; plasmid; vaccine; plague; ds.  
OS Yersinia pestis.

EH Key Location/Qualifiers  
FT CDS 17..532  
FT /\*tag= a  
FT /product= "F1 antigen"  
PN W09824912-A2.  
PD 11-JUN-1998  
PF 04-DEC-1997: U22617.  
PR 04-DEC-1996: US-767115.  
PA (HESK-) HESKA CORP.  
PI Haanes EJ, Osorio JE, Thomas RE;  
DR WPI: 98-33331/29.  
DR P-PSDB: W59783.  
PT Protection of animals against plague - using nucleic acid encoding  
PT epitopes from Yersinia, Pasteurella and Francisella spp.  
PS Claim 8; Pages 53-54; 75pp; English.  
CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used  
CC in the method of the invention. Plasmid and host cells are used to  
CC produce recombinant antigens, especially Yersinia pestis antigens.  
CC The recombinant antigens can be used in vaccines that are capable of  
CC protecting an animal from contracting plague.  
SQ Sequence 544 BP; 166 A; 118 C; 112 G; 148 T;

Query Match 33.0%; Score 516.8; DB 1; Length 544;  
Best Local Similarity 99.6%; Pred. No. 1.8e-128;  
Matches 518; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 68 ATATGAAAAAATCAGTCCGCTTATCGCCATTGATTTGTAACCTTATGCACTGCTA 127  
DB 15 ATATGAAAAAATCAGTCCGCTTATCGCCATTGATTTGTAACCTTATGCACTGCTA 74  
QY 128 ATGCGGAGATTTAATGCAAGCACACCTGCAAGCGCACTCTGTTGAACAGCCGCA 187  
DB 75 ATGCGGAGATTTAATGCAAGCACACCTGCAAGCGCACTCTGTTGAACAGCCGCA 134  
QY 188 TCACCTTACATATATAGGAAGGCGCTCCATTTACATTTATGCAATATGAAACATCGATA 247  
DB 135 TCACCTTACATATATAGGAAGGCGCTCCATTTACATTTATGCAATATGAAACATCGATA 194  
QY 248 CAGATTACTGTTGGTAGCGCTTACTCTTGCGGCTATTAACAAGAACCACTAGACAT 307  
DB 195 CAGATTACTGTTGGTAGCGCTTACTCTTGCGGCTATTAACAAGAACCACTAGACAT 254  
QY 308 CTGTTAACTTTACAGATGCCGCGGGTGATCCCATGTACTTAACATTTACTTCAGATG 367  
DB 255 CTGTTAACTTTACAGATGCCGCGGGTGATCCCATGTACTTAACATTTACTTCAGATG 314  
QY 368 GAAATTAACCAACCAATTTACTACAAAGATGATGGCAAGGATTTAGAGATTTTGTATCT 427  
DB 315 GAAATTAACCAACCAATTTACTACAAAGATGATGGCAAGGATTTAGAGATTTTGTATCT 374  
QY 428 CTCCTAAGTAAACGGTAGAACCTTGAGGGGATGACGTGCTTGCTAGGGGCAACC 487  
DB 375 CTCCTAAGTAAACGGTAGAACCTTGAGGGGATGACGTGCTTGCTAGGGGCAACC 434  
QY 488 AGGATTTCTTTGTCGCTCAATTTGTTCCAAAGCGGTAAACTTGACAGCAGTAAATACA 547  
DB 435 AGGATTTCTTTGTCGCTCAATTTGTTCCAAAGCGGTAAACTTGACAGCAGTAAATACA 494  
QY 548 CTGATGCTGTAAACCGTATCTATCTAACAAGATTTCAT 587  
DB 495 CTGATGCTGTAAACCGTATCTATCTAACAAGATTTCAT 534

RESULT 11  
V41595  
ID V41595 standard; DNA; 510 BP.  
AC V41595;  
DT 26-OCT-1998 (first entry)  
DE Nucleotide sequence of F1 antigen nypf1(b)sec544.  
KW F1 antigen; plasmid; vaccine; plague; ds.  
OS Yersinia pestis.  
PN W09824912-A2.

```
PD 11-JUN-1998.
PR 04-DEC-1997; U22617.
PR 04-DEC-1996; US-767115.
PA (HESK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR WPI: 98-333331/29.
PT Protection of animals against plague - using nucleic acid encoding
PT antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 8; Page 53; 75pp; English.
CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
SQ Sequence 510 BP; 153 A; 114 C; 106 G; 137 T;

Query Match 32.6%; Score 510; DB 1; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.1e-16;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 ATGAAAAAATAGTCCCTTATCGCATTTGATTTGAACTATTCGAACGTCTAT 129
DB 1 ATGAAAAAATAGTCCCTTATCGCATTTGATTTGAACTATTCGAACGTCTAT 60
OY 130 GCGGAGATTACTGCAAGACACACTGCAAGCGCACTCTTTGACAGCCGCATC 189
DB 61 GCGGAGATTACTGCAAGACACACTGCAAGCGCACTCTTTGACAGCCGCATC 120
OY 190 ACTCTTACATATAGAGAGCGCGCTCCATTTACATTTATGACANATGAAACATCAT 249
DB 121 ACTCTTACATATAGAGAGCGCGCTCCATTTACATTTATGACANATGAAACATCAT 180
OY 250 GAATTACTTGTGATGATGCTTACTCTTGGCGGTATTAAGAGCACTGACATCT 309
DB 181 GAATTACTTGTGATGATGCTTACTCTTGGCGGTATTAAGAGCACTGACATCT 240
OY 310 GTTAATCTTACAGATGCCCGGGTGATCCCATGACTTAACTTACTCTCAGATGA 369
DB 241 GTTAATCTTACAGATGCCCGGGTGATCCCATGACTTAACTTACTCTCAGATGA 300
OY 370 AATAACCAACCAATTCATCAAAAAGTATGGCAGAGATTCTAGAGATTTGATATCTCT 429
DB 301 AATAACCAACCAATTCATCAAAAAGTATGGCAGAGATTCTAGAGATTTGATATCTCT 360
OY 430 CCTAAGGTAAAGGTGAGAACTTGTGGGGGATGACGTGCTTGGCTACGGGACCCAG 489
DB 361 CCTAAGGTAAAGGTGAGAACTTGTGGGGGATGACGTGCTTGGCTACGGGACCCAG 420
OY 490 GATTTCCTTGTGCTCAATTTGGTCCAAAGGGGTAACCTTGACAGGTAATACACT 549
DB 421 GATTTCCTTGTGCTCAATTTGGTCCAAAGGGGTAACCTTGACAGGTAATACACT 480
OY 550 GATGCTGAACCGTAACCGTATCTAACCA 579
DB 481 GATGCTGAACCGTAACCGTATCTAACCA 510

RESULT 12
ID V41600 standard; DNA; 474 BP.
AC V41600:
DT 26-OCT-1998 (first entry)
DE Nucleotide sequence of F1 antigen nypF1mat474.
KW F1 antigen; plasmid; vaccine; plaque; ds.
OS Yersinia pestis.
FH Key Location/Qualifiers
FT CDS 7..459
FT /tag= a
FT /product= "F1 antigen"
PN W09824912-A2.
PD 11-JUN-1998.
PF 04-DEC-1997; U22617.
```

```
PR 04-DEC-1996; US-767115.
PA (HESK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR WPI: 98-333331/29.
DR P-PSDB: W59787.
PT Protection of animals against plague - using nucleic acid encoding
PT antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 8; Pages 59-60; 75pp; English.
CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
SQ Sequence 474 BP; 143 A; 106 C; 99 G; 126 T;

Query Match 28.9%; Score 452.8; DB 1; Length 474;
Best Local Similarity 99.6%; Pred. No. 1.9e-11;
Matches 454; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 132 GCGAGATTAACTGCAAGACACCTGCAAGCGCACTCTTGTGAACCGCCGCATCAC 191
DB 9 GCGAGATTAACTGCAAGACACCTGCAAGCGCACTCTTGTGAACCGCCGCATCAC 68
OY 192 TCTTACATATAGAGAGCGGCTCCATTTACATTTATGACAAATGGAACATGATACAGA 251
DB 69 TCTTACATATAGAGAGCGGCTCCATTTACATTTATGACAAATGGAACATGATACAGA 128
OY 252 ATTACTTGTGTGATGCTTACTCTTGGCGGCTATTAACAGAGACCACTAGCATCTGT 311
DB 129 ATTACTTGTGTGATGCTTACTCTTGGCGGCTATTAACAGAGACCACTAGCATCTGT 188
OY 312 TAACCTTACAGATGCCGGGGTGATCCCATGACTTAACTTACTCTCAGATGGAAA 371
DB 189 TAACCTTACAGATGCCGGGGTGATCCCATGACTTAACTTACTCTCAGATGGAAA 248
OY 372 TAACCAACCAATTCATCAAAAAGTATGGCAAGATTTCTAGATTTGATATCTCTCC 431
DB 249 TAACCAACCAATTCATCAAAAAGTATGGCAAGATTTCTAGATTTGATATCTCTCC 308
OY 432 TAAGTAAAGGCTGAGAACTTGTGGGGGATGACGTGCTTGGCTACGGGACCCAGA 491
DB 309 TAAGTAAAGGCTGAGAACTTGTGGGGGATGACGTGCTTGGCTACGGGACCCAGA 368
OY 492 TTTCTTTTTCCTCAATTTGTTCCAAAGCGGCTAACTTGACAGAGTAATACACTGA 551
DB 369 TTTCTTTTTCCTCAATTTGTTCCAAAGCGGCTAACTTGACAGAGTAATACACTGA 428
OY 552 TGCTGAACCGTAACCGTATCTAACCAAGATTCAT 587
DB 429 TGCTGAACCGTAACCGTATCTAACCAAGATTCAT 464

RESULT 13
ID 092817 standard; DNA; 541 BP.
AC 092817:
DT 17-DEC-1995 (first entry)
DE Yersinia pestis cafI (F1) antigen in plasmid pFGAL2a.
KW Vaccine; antigen; Salmonella typhimurium; Salmonella typhi;
KW bubonic plague; pneumonic plague; ds.
OS Yersinia pestis.
FH Key Location/Qualifiers
FT CDS 2..454
FT /tag= a
FT /tag= a
FT misc_feature 1..6
FT /note= b
FT /note= "lacZ promoter fusion site"
FT FT 536..541
FT /tag= c
FT /note= "vector pFGAL2a bases"
PN W09518231-A1.
PD 06-JUL-1995.
```

PF 23-DEC-1994: G02818.  
PR 24-DEC-1993: GB-026425.  
PA (MINA ) UK SEC FOR DEFENCE.  
PI Howells A, Leary SEC, Oyston PCF, Tibball RW, Williamson ED;  
DR WPI: 95-246396/32.  
P-PSDB: R76526.  
PT DNA constructs capable of transforming microorganisms - which can be  
used as live or attenuated vaccines which induce an immune response,  
PT against Yersinia pestis, at mucosal surfaces.  
PS Claim 7: Page 15-16; 27pp; English.  
CC The sequence represents the plasmid pFGAL2a construct showing the  
fusion of the first few bases of beta-galactosidase in the vector  
CC with the Y. pestis cafi (F1) antigen minus its signal sequence and  
having a 5' tail including a SacI restriction site, and up to the  
CC cafi AAC-3' end with some vector bases. The DNA construct can be  
used to transform human or animal gut colonizing microorganisms,  
CC specifically attenuated Salmonella typhimurium or Salmonella typh.  
CC The transformed microorganisms can be used as live/attenuated  
CC vaccines which induce immune responses at mucosal surfaces. The  
CC vaccines provide protection against infection with Y. pestis, and  
CC are parenterally and orally active vaccines offering protection  
CC against bubonic and pneumonic plague.  
SQ Sequence 541 BP; 163 A; 120 C; 111 G; 147 T;  
  
Query Match 28.9%; Score 452.2; DB 1; Length 541;  
Best Local Similarity 99.3%; Pred. No. 2.8e-111;  
Matches 454; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 131 CGGCAGATTAACTGACAGCACCAGCACTGTCCTGTTGAACGCCCGCATCA 190  
DB 6 CGGCAGATTAACTGACAGCACCAGCACTGTTGTTGAACGCCCGCATCA 65  
  
QY 191 CTCTACATTAAGAGGAGGCGCTCCATTACATTTATGACAATGAAATCGATACAG 250  
DB 66 CTATTTACATTAAGAGGAGGCGCTCCATTACATTTATGACAATGAAATCGATACAG 125  
  
QY 251 AATTACTTGTGGTACGCTTACTCTTGGCGGTATAAAACAGAACCATGACATCTG 310  
DB 126 AATTACTTGTGGTACGCTTACTCTTGGCGGTATAAAACAGAACCATGACATCTG 185  
  
QY 311 TTAACTTTAGAGTGGCGGGGATGCCATGACTTAACTTAACTTCTCAGATGGA 370  
DB 186 TTAACTTTAGAGTGGCGGGGATGCCATGACTTAACTTAACTTCTCAGATGGA 245  
  
QY 371 ATACACCAATTCACATACAAAGATGTTGGCAAGATTTAGAGATTTGATCTCTC 430  
DB 246 ATACACCAATTCACATACAAAGATGTTGGCAAGATTTGATGATTTGATCTCTC 305  
  
QY 431 CTAAAGTAAACGGTGAGAACCTTGTGGGGATGACGTGCTTGGCTACGGGACGAG 490  
DB 306 CTAAAGTAAACGGTGAGAACCTTGTGGGGATGACGTGCTTGGCTACGGGACGAG 365  
  
QY 491 ATTCTTTGTCGCTCAATTGTTCCAAAGCGGTAACTTGACAGAGTAATACCTG 550  
DB 366 ATTCTTTGTCGCTCAATTGTTCCAAAGCGGTAACTTGACAGAGTAATACCTG 425  
  
QY 551 ATGCTGAACCGTAAACCGTATCTAACCAAGATTTCAT 587  
DB 426 ATGCTGAACCGTAAACCGTATCTAACCAATATTCAT 462  
  
RESULT 14  
ID T38244  
T38244 standard: DNA; 541 BP.  
AC T38244;  
DT 28-DEC-1996 (first entry)  
DE Y. pestis cafi (F1 antigen) gene.  
KW Plaque; vaccine; genetic immunisation; F1 antigen; cafi;  
V antigen; ds.  
OS Yersinia pestis strain GB.  
FH Key misc\_recomb Location/Qualifiers  
FT misc\_recomb 1.6

FT FT /tag- a  
FT FT /note- "Includes the first few bases of the  
FT FT beta-galactosidase sequence in vector  
FT FT pFGAL2a"  
FT cds 2..457  
FT FT /tag- b  
FT FT /product- mature F1 antigen  
FT FT 536..541  
FT misc\_recomb /tag- c  
PN W09628551-A1.  
PD 19-SEP-1996.  
PR 13-MAR-1995: G00571.  
PR 13-MAR-1995: GB-005059.  
PR 15-SEP-1995: GB-018946.  
PR 05-DEC-1995: GB-024825.  
PA (MINA ) UK SEC FOR DEFENCE.  
PI Bennett AM, Leary SEC, Oyston PCF, Tibball RW, Williamson ED;  
DR WPI: 96-433824/43.  
P-PSDB: W01042.  
PT Yersinia pestis V antigen and F1 antigen or their protective  
PT epitopic parts - useful in vaccine for protection against plague  
PS Claim 7: Page 43-45; 98pp; English.  
CC The cafi gene (T38244) codes for the Yersinia pestis F1 antigen  
CC (W01042), which is capable of evoking protective immune responses  
CC in animals. The gene was amplified from Y. pestis DNA by PCR using  
CC primers (T38245-46) homologous to the 5' and 3' ends of the gene.  
CC The cafi PCR product was ligated into pUC18 and transformed into E.  
CC coli JM109 to produce vector pFGAL2a. The cafi gene can be used to  
CC produce recombinant F1 antigen for use in vaccines against plague.  
CC Expression in gut-colonising organisms and attenuated Salmonella  
CC typh allows prodn. of live vaccines. The gene can itself be used  
CC in genetic vaccines. F1/V antigen fusions were also created (see  
CC also T38249 and T38256).  
SQ Sequence 541 BP; 163 A; 120 C; 111 G; 147 T;  
  
Query Match 28.9%; Score 452.2; DB 1; Length 541;  
Best Local Similarity 99.3%; Pred. No. 2.8e-111;  
Matches 454; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 131 CGGCAGATTAACTGACAGCACCAGCACTGTCCTGTTGAACGCCCGCATCA 190  
DB 6 CGGCAGATTAACTGACAGCACCAGCACTGTTGTTGAACGCCCGCATCA 65  
  
QY 191 CTCTACATTAAGAGGAGGCGCTCCATTACATTTATGACAATGAAATCGATACAG 250  
DB 66 CTATTTACATTAAGAGGAGGCGCTCCATTACATTTATGACAATGAAATCGATACAG 125  
  
QY 251 AATTACTTGTGGTACGCTTACTCTTGGCGGTATAAAACAGAACCATGACATCTG 310  
DB 126 AATTACTTGTGGTACGCTTACTCTTGGCGGTATAAAACAGAACCATGACATCTG 185  
  
QY 311 TTAACTTTAGAGTGGCGGGGATGCCATGACTTAACTTAACTTCTCAGATGGA 370  
DB 186 TTAACTTTAGAGTGGCGGGGATGCCATGACTTAACTTAACTTCTCAGATGGA 245  
  
QY 371 ATACACCAATTCACATACAAAGATGTTGGCAAGATTTAGAGATTTGATCTCTC 430  
DB 246 ATACACCAATTCACATACAAAGATGTTGGCAAGATTTGATGATTTGATCTCTC 305  
  
QY 431 CTAAAGTAAACGGTGAGAACCTTGTGGGGATGACGTGCTTGGCTACGGGACGAG 490  
DB 306 CTAAAGTAAACGGTGAGAACCTTGTGGGGATGACGTGCTTGGCTACGGGACGAG 365  
  
QY 491 ATTCTTTGTCGCTCAATTGTTCCAAAGCGGTAACTTGACAGAGTAATACCTG 550  
DB 366 ATTCTTTGTCGCTCAATTGTTCCAAAGCGGTAACTTGACAGAGTAATACCTG 425  
  
QY 551 ATGCTGAACCGTAAACCGTATCTAACCAAGATTTCAT 587  
DB 426 ATGCTGAACCGTAAACCGTATCTAACCAATATTCAT 462

RESULT 15

092818 092818 standard: DNA; 542 BP.

AC 092818;

DE Yersinia pestis caf1 (F1) antigen in plasmid pF51G3a.

KW vaccine; antigen; Salmonella typhimurium; Salmonella typhi;

OS Yersinia pestis.

FT Key Location/Qualifiers

FT cds 3..455

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

FT misc-feature 1..6

QY 489 GGATTTCTTTGTTGCTCAATTGGTTCCAAAGGCGGTAAACTTGACAGAGTAATAACAC 548  
 Db 365 GGATTTCTTTGTTGCTCAATTGGTTCCAAAGGCGGTAAACTTGACAGAGTAATAACAC 424  
 QY 549 TGATGCTGTACCGGTACCGGTATCTTACCAAGAAATTCAT 587  
 Db 425 TGATGCTGTACCGGTACCGGTATCTTACCAAGAAATTCAT 463

Search completed: November 27, 1999, 19:34:03  
 Job time: 2661 sec

Query Match 28.8%; Score 451; DB 1; Length 542;  
 Best Local Similarity 98.9%; Pred. No. 5.9e-111;  
 Matches 454; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 129 TCGGCGAGATTACTACTGACGACACTGCAACGCAACTTTGTAACGAGCCGCAT 188  
 Db 5 TCCCGCGAGATTACTACTGACGACACTGCAACGCAACTTTGTAACGAGCCGCAT 64  
 QY 189 CACTTTACATATAAGGAGGCGCTCCAAATTAATATGAGCAATGGAACATCGATAC 248  
 Db 65 CACTTTACATATAAGGAGGCGCTCCAAATTAATATGAGCAATGGAACATCGATAC 124  
 QY 249 AGAATTTACTTTGTTGCTTACTTCTTGGCGGTATATAAGGAGCACTAGACATC 308  
 Db 125 AGAATTTACTTTGTTGCTTACTTCTTGGCGGTATATAAGGAGCACTAGACATC 184  
 QY 309 TGTAACTTACAGATCCCGGGGTATCCCATGTACTTAACATTTCTCTCGAGATGG 368  
 Db 185 TGTAACTTACAGATCCCGGGGTATCCCATGTACTTAACATTTCTCTCGAGATGG 244  
 QY 369 AAATAACCAACCAATTCCTACCAAAAGTATGGCAAGATTCAGAGATTTGATATCTC 428  
 Db 245 AAATAACCAACCAATTCCTACCAAAAGTATGGCAAGATTCAGAGATTTGATATCTC 304  
 QY 429 TCCTAAGGTAAAGGTAGAACTTTGGGGGATGACGTCCTTGGCTACGGGACGCA 488  
 Db 305 TCCTAAGGTAAAGGTAGAACTTTGGGGGATGACGTCCTTGGCTACGGGACGCA 364

**THIS PAGE BLANK (USPTO)**





Percent Similarity: 100.000 Percent Identity: 94.444

alignment\_block:  
US-08-699-716A-2 x US-08-480-604A-25 (1) ..

Align seg 1/1 to: US-08-480-604A-25 from: 1 to: 1402

1 MGHNNNNNNHSSGHID 18  
|||||  
1 MGHNNNNNNHSSGHIE 18

seq\_name: /cgn2\_6/ptodata/1/ina/5D\_COMB.seq:US-08-405-496A-25

seq\_documentation\_block:  
; Sequence 25, Application US/08405496A  
; Patent No. 5919665  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, JAMES A.  
; TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM  
; NUMBER OF INVENTION: NEUROTOXIN  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: MEDLEN & CARROLL, LLP  
; STREET: 220 MONTGOMERY STREET, SUITE 2200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/405,496A  
; FILING DATE: 16-MAR-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/329,154  
; FILING DATE: 25-OCT-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/161,907  
; FILING DATE: 02-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/985,321  
; FILING DATE: 04-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/429,791  
; FILING DATE: 31-OCT-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: INGOLIA, DIANE E.  
; REGISTRATION NUMBER: 40,027  
; REFERENCE/DOCKET NUMBER: OPHD-01308  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1402 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1386  
; US-08-405-496A-25

alignment\_scores:  
Quality: 93.00 Length: 18  
Ratio: 5.167 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 94.444

alignment\_block:  
US-08-699-716A-2 x US-08-405-496A-25 (1) ..

Align seg 1/1 to: US-08-405-496A-25 from: 1 to: 1402

1 MGHNNNNNNHSSGHID 18  
|||||  
1 MGHNNNNNNHSSGHIE 18

seq\_name: /cgn2\_6/ptodata/1/ina/5C\_COMB.seq:US-08-332-766A-1

seq\_documentation\_block:  
; Sequence 1, Application US/08332766A  
; Patent No. 5843647  
; GENERAL INFORMATION:  
; APPLICANT: JEFFREYS, Alec J.  
; APPLICANT: ARMOUR, John  
; TITLE OF INVENTION: SIMPLE TANDEM REPEATS  
; NUMBER OF SEQUENCES: 125  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.  
; STREET: 1100 New York Avenue, N.W.  
; CITY: Washington  
; STATE: D. C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-3918  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/332,766A  
; FILING DATE: 01-NOV-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9326052.9  
; FILING DATE: 21-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BIRD, Donald J.  
; REGISTRATION NUMBER: 25,323  
; REFERENCE/DOCKET NUMBER: 217211/M94/0434/GB  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 861-3000  
; TELEFAX: (202) 822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 377 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-08-332-766A-1

alignment\_scores:  
Quality: 69.00 Length: 20  
Ratio: 4.059 Gaps: 0  
Percent Similarity: 85.000 Percent Identity: 55.000

alignment\_block:  
US-08-699-716A-2 x US-08-332-766A-1/rev (5) ..

Align seg 1/1 to reverse of: US-08-332-766A-1 from: 1 to: 377

3 NNNNNNNHSSGHIDDDK 22  
||||| : : : : :  
37 NNNNNNNHSSGHIDDDK 18

seq\_name: /cgn2\_6/ptodata/1/ina/5D\_COMB.seq:US-08-600-783-15

```

seq_documentation_block:
; Sequence 15, Application US/08600783
; Patent No. 5962267
;
; GENERAL INFORMATION:
; APPLICANT: SHIN, Hang Cheol
; APPLICANT: CHANG, Seung Gu
; APPLICANT: KIM, Chong Suh1
; TITLE OF INVENTION: Proinsulin Derivative and Process
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SHIN, Hang Cheol
; STREET: Seangma-Hanshin Apt. 102-1206,
; STREET: #245 Cholsan-dong
; CITY: Kwangmyung-shi
; STATE: Kyungki-do
; COUNTRY: Republic of Korea
; ZIP: 423-030
; ADDRESSEE: CHANG, Seung Gu
; STREET: Hyundai Apt. 71-203, Apkujong-dong,
; STREET: Kangnam-ku
; CITY: Seoul
; STATE: Seoul
; COUNTRY: Republic of Korea
; ZIP: 135-110
; ADDRESSEE: KIM, Dae Young
; STREET: Sosa Jukong Apt. 108-202, Sosa Bon-dong,
; STREET: Sosa-ku
; CITY: Bucheon-shi
; STATE: Kyungki-do
; COUNTRY: Republic of Korea
; ZIP: 422-230
; ADDRESSEE: KIM, Chong Suh1
; STREET: Garden Heights Apt. 202-801, #100,
; STREET: Hwangkeum-dong, Soosung-ku
; CITY: Taegu
; STATE: Taegu
; COUNTRY: Republic of Korea
; ZIP: 706-040
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk, 3.5 inch, 1.44MB storage
; COMPUTER: IBM PC/AT
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600,783
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: KR 95-2751
; FILING DATE: 15-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Shahan Islam
; REGISTRATION NUMBER: 32,507
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-1000
; TELEFAX: (212) 953-7249
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: primer DNA
; ANTI-SENSE: no
; US-08-600-783-15

```

```

alignment_scores:
Quality: 62.00      Length: 11
Ratio: 5.636       Gaps: 0

```

```

Percent Similarity: 100.000   Percent Identity: 100.000
alignment_block:
US-08-699-716a-2 x US-08-600-783-15 (1)
Align seg 1/1 to: US-08-600-783-15 from: 1 to: 47
3 HNNNNNNHNS 13
|||||
5 HNNNNNNHNS 15

```

```

seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-068-747-8
seq_documentation_block:
; Sequence 8, Application US/08068747
; Patent No. 5695933
;
; GENERAL INFORMATION:
; APPLICANT: Schalling, Martin
; APPLICANT: Hudson, Thomas J.
; APPLICANT: Housman, David E.
; TITLE OF INVENTION: Direct Determination of Expanded
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millita Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/068,747
; FILING DATE: 28-MAY-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: MIT-6141
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic"
; US-08-068-747-8

```

```

alignment_scores:
Quality: 60.00      Length: 10
Ratio: 6.000       Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

```

```

alignment_block:
US-08-699-716a-2 x US-08-068-747-8/rev (5)
Align seg 1/1 to reverse of: US-08-068-747-8 from: 1 to: 36
3 HNNNNNNHNS 12
|||||
11 HNNNNNNHNS 2

```

```

seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-563-864-1

```

```
seq_documentation_block:
; Sequence 1, Application US/08563864
; Patent No. 5674687
; GENERAL INFORMATION:
; APPLICANT: Hershfield, Bennett
; TITLE OF INVENTION: A METHOD FOR THE IDENTIFICATION OF
; TITLE OF INVENTION: NUCLEIC ACID SAMPLES FROM DNA-CONTAINING ORGANISMS
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/563,864
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/251,395
; FILING DATE: 31-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman Esq., Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/160 (D-1452)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; US-08-563-864-1

alignment_scores:
; Quality: 60.00 Length: 10
; Ratio: 6.000 Gaps: 0
; Percent Similarity: 100.000 Percent Identity: 100.000

seq_documentation_block:
; Sequence 1, Application PC/TUS9506839
; GENERAL INFORMATION:
; APPLICANT: Hershfield, Bennett
; TITLE OF INVENTION: A METHOD FOR THE IDENTIFICATION
; TITLE OF INVENTION: OF NUCLEIC ACID SAMPLES FROM DNA-CONTAINING ORGANISMS
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/781,034
; FILING DATE: 19911018
```

```
STATE: New York
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06839
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/251,395
FILING DATE: 31-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Goldman Esq., Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/160 (D-1452)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHEICAL: NO
ANTI-SENSE: NO
PCT-US95-06839-1
```

```
alignment_scores:
; Quality: 60.00 Length: 10
; Ratio: 6.000 Gaps: 0
; Percent Similarity: 100.000 Percent Identity: 100.000
```

```
seq_documentation_block:
; Sequence 4, Application US/07781034
; Patent No. 5442050
; GENERAL INFORMATION:
; APPLICANT: Fishman, Jay A.
; TITLE OF INVENTION: Molecular Cloning of Antigens Shared By
; TITLE OF INVENTION: Rat- and Human-Derived Pneumocystis Carinii
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/781,034
; FILING DATE: 19911018
```

```
Align seg 1/1 to: PCT-US95-06839-1 from: 1 to: 38
3 HNNNNNNH 12
|||||
2 HNNNNNNH 11
```

```
seq_name: /cgn2_6/ptodata/1/ina/5A.COMB.seq:US-07-781-034-4
```

```
seq_documentation_block:
; Sequence 4, Application US/07781034
; Patent No. 5442050
; GENERAL INFORMATION:
; APPLICANT: Fishman, Jay A.
; TITLE OF INVENTION: Molecular Cloning of Antigens Shared By
; TITLE OF INVENTION: Rat- and Human-Derived Pneumocystis Carinii
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/781,034
; FILING DATE: 19911018
```

```

; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
;   NAME: Granahan, Patricia
;   REGISTRATION NUMBER: 32,227
;   REFERENCE/DOCKET NUMBER: MGH91-02A
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (617) 861-6240
;     TELEFAX: (617) 861-9540
;   INFORMATION FOR SEQ ID NO: 4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 1189 base pairs
;       TYPE: NUCLEIC ACID
;       STRANDEDNESS: double
;       TOPOLOGY: linear
;     MOLECULE TYPE: cDNA to mRNA
;     HYPOTHETICAL: NO
;     ANTI-SENSE: NO
;   US-07-781-034-4

alignment_scores:
  Quality: 70.00      Length: 15
  Ratio: 5.385      Gaps: 0
  Percent Similarity: 86.667      Percent Identity: 80.000

alignment_block:
  US-08-699-716A-2 x US-07-781-034-4/rev (4)  ..

Align seg 1/1 to reverse of: US-07-781-034-4 from: 1 to: 1189
  3 HHHHHHHHHSSGHT 17
    |||||
  203 HHHHHHHHHSHI 189

seq_name: /cgn2_6/prodata/1/lna/PCTUS9_COMB.seq:PCT-US92-08328-4

seq_documentation_block:
; Sequence 4, Application PC/TUS9208328
; GENERAL INFORMATION:
;   APPLICANT: Fishman, Jay A.
;   TITLE OF INVENTION: Molecular Cloning of Antigens Shared By
;   TITLE OF INVENTION: Rat- and Human-Derived Pneumocystis Carinii
;   NUMBER OF SEQUENCES: 4
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
;     STREET: Two Militia Drive
;     CITY: Lexington
;     STATE: Massachusetts
;     COUNTRY: USA
;     ZIP: 02173
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: Patentin Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: PCT/US92/08328
;     FILING DATE: 19920930
;   CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: US 07/781,034
;       FILING DATE: 18-OCT-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: US 07/768,166
;       FILING DATE: 30-SEP-1991
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Granahan, Patricia
;     REGISTRATION NUMBER: 32,227
;     REFERENCE/DOCKET NUMBER: MGH91-02AA PCT
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (617) 861-6240
;       TELEFAX: (617) 861-9540
;       TELEX: 951794
```

```

; INFORMATION FOR SEQ ID NO: 4:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1189 base pairs
;     TYPE: NUCLEIC ACID
;     STRANDEDNESS: double
;     TOPOLOGY: linear
;   MOLECULE TYPE: cDNA to mRNA
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   PCT-US92-08328-4

alignment_scores:
  Quality: 70.00      Length: 15
  Ratio: 5.385      Gaps: 0
  Percent Similarity: 86.667      Percent Identity: 80.000

alignment_block:
  US-08-699-716A-2 x PCT-US92-08328-4/rev (4)  ..

Align seg 1/1 to reverse of: PCT-US92-08328-4 from: 1 to: 1189
  3 HHHHHHHHHSSGHT 17
    |||||
  203 HHHHHHHHHSHI 189

seq_name: /cgn2_6/prodata/1/lna/5C_COMB.seq:US-08-332-766A-1

seq_documentation_block:
; Sequence 1, Application US/08332766A
; Patent No. 5843647
; GENERAL INFORMATION:
;   APPLICANT: JEFFREYS, Alec J.
;   TITLE OF INVENTION: SIMPLE TANDEM REPEATS
;   NUMBER OF SEQUENCES: 125
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
;     STREET: 1100 New York Avenue, N.W.
;     CITY: Washington
;     STATE: D.C.
;     COUNTRY: U.S.A.
;     ZIP: 20005-3918
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     OPERATING SYSTEM: IBM PC compatible
;     SOFTWARE: Patentin Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/332,766A
;     FILING DATE: 01-NOV-1994
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: GB 9326052.9
;     FILING DATE: 21-DEC-1993
;   ATTORNEY/AGENT INFORMATION:
;     NAME: BIRD, Donald J.
;     REGISTRATION NUMBER: 25,323
;     REFERENCE/DOCKET NUMBER: 217211/M94/0434/CB
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (202) 861-3000
;       TELEFAX: (202) 822-0944
;       TELEX: 6714627 CUSH
;   INFORMATION FOR SEQ ID NO: 1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 377 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;     MOLECULE TYPE: DNA (genomic)
;   US-08-332-766A-1
```

US-08-699-716A-2 x US-08-332-766A-30/rev (6)

APPLICANT: Plaut, Andrew G.

APPLICANT: Wright, Andrew  
TITLE OF INVENTION: HELICOBACTER PYLORI NICKEL  
TITLE OF INVENTION: BINDING PROTEIN  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/05772  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul C.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00398/090001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 183 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
PCT-US95-05772-3

alignment\_scores:  
Quality: 60.00 Length: 22  
Ratio: 3.750 Gaps: 0  
Percent Similarity: 72.727 Percent Identity: 54.545

alignment\_block:  
US-08-699-716A-2 x PCT-US95-05772-3 (1) ..  
Align seg 1/1 to: PCT-US95-05772-3 from: 1 to: 183

2 GHHNNHHNNHSSGHIDDDKH 23  
||||||| | : : : : |  
10 GHHNNHHNNHNNHYGGEHHN 31

seq\_name: /cgn2\_6/ptodata/1/ina/5C\_COMB.seq:US-08-203-532F-3

seq\_documentation\_block:  
Sequence 3, Application US/08203532F  
Patent No. 5856121  
GENERAL INFORMATION:  
APPLICANT: Gorski, David H.  
APPLICANT: Walsh, Kenneth  
TITLE OF INVENTION: Growth Arrest Homeobox Gene  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Calfee, Halter, and Griswold  
STREET: 800 Superior Avenue  
CITY: Cleveland  
STATE: Ohio  
COUNTRY: U.S.A.  
ZIP: 44114-2688  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/203,532F  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Golrick, Mary E.  
REGISTRATION NUMBER: 34829  
REFERENCE/DOCKET NUMBER: 22311/00114  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (216) 622-8200  
TELEFAX: (216) 241-0816  
TELEX: 980499  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 941 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 33..941  
US-08-203-532F-3

alignment\_scores:  
Quality: 64.00 Length: 11  
Ratio: 5.818 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-08-699-716A-2 x US-08-203-532F-3 (3) ..  
Align seg 1/1 to: US-08-203-532F-3 from: 1 to: 941

2 GHHNNHHNNH 12  
||||||| |  
77 GHHNNHHNNH 87

seq\_name: /cgn2\_6/ptodata/1/ina/PCTUS9\_COMB.seq:PCT-US95-01882A-3

seq\_documentation\_block:  
Sequence 3, Application PC/TUS9501882A  
GENERAL INFORMATION:  
APPLICANT: Gorski, David H.  
APPLICANT: Walsh, Kenneth  
TITLE OF INVENTION: Growth Arrest Homeobox Gene  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Calfee, Halter, and Griswold  
STREET: 800 Superior Avenue  
CITY: Cleveland  
STATE: Ohio  
COUNTRY: U.S.A.  
ZIP: 44114-2688  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/01882A  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Golrick, Mary E.  
REGISTRATION NUMBER: 34829  
REFERENCE/DOCKET NUMBER: 22311/00114  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (216) 622-8200

TELEFAX: (216) 241-0816  
 TELEX: 980499  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 941 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: both  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 33..941  
 PCT-US95-01882A-3

alignment\_scores:  
 Quality: 64.00 Length: 11  
 Ratio: 5.818 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-699-716A-2 x PCT-US95-01882A-3 (3)

Align seg 1/1 to: PCT-US95-01882A-3 from: 1 to: 941

2 GHHHHHHHHH 12  
 1111111111  
 77 GHHHHHHHHH 87



GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 1999, 17:36:42 ; Search time 27.29 Seconds  
(without alignments)  
6068.421 Million cell updates/sec

Title: US-08-699-716a-1

Perfect score: 1566  
Sequence: 1 ATGGGCCATCATCATCATCA.....ATGACACGTCTGTAAATGA 1566

Scoring table: IDENTITY\_MNC

Searched: 195662 seqs, 52875861 residues

Database : Issued\_Patents\_NA:\*

1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/5C.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/5D.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCITUS9.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/Dackfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	53.4	3.4	1402	2	US-08-480-604A-25	Sequence 25, Appl
2	53.4	3.4	1402	4	US-08-405-496A-25	Sequence 25, Appl
3	48	3.1	7218	2	US-08-232-463A-14	Sequence 14, Appl
4	39	2.5	4435	2	US-08-308-872B-5	Sequence 5, Appl
5	37.4	2.4	4365	5	PCT-US91-08525-31	Sequence 31, Appl
6	37.4	2.4	4365	5	PCT-US93-04384-11	Sequence 11, Appl
7	35	2.2	47	4	US-08-600-783-15	Sequence 15, Appl
8	34.8	2.2	2415	4	US-08-785-428-1	Sequence 1, Appl
9	34.8	2.2	1446	4	US-08-560-007B-4	Sequence 4, Appl
10	34.6	2.2	1100	3	US-08-179-557-20	Sequence 20, Appl
11	34	2.2	867	4	US-09-036-582-36	Sequence 36, Appl
12	33.2	2.1	10564	1	US-08-206-176-5	Sequence 5, Appl
13	33.2	2.1	9636	2	US-08-323-170B-1	Sequence 1, Appl
14	33	2.1	1888	2	US-08-229-145-13	Sequence 13, Appl
15	33	2.1	8920	3	US-08-446-855A-1	Sequence 1, Appl
16	33	2.1	2993	4	US-08-415-593-42	Sequence 42, Appl
17	32.8	2.1	891	3	US-07-941-523-19	Sequence 19, Appl
18	32.8	2.1	1959	3	US-08-137-175A-1	Sequence 1, Appl
19	32.8	2.1	3369	4	US-08-525-742-1	Sequence 6, Appl
20	32.6	2.1	3311	1	US-07-777-715-6	Sequence 8, Appl
21	32.6	2.1	2799	1	US-07-777-715-8	Sequence 8, Appl
22	32.6	2.1	2799	1	US-07-960-932-1	Sequence 1, Appl
23	32.6	2.1	2838	1	US-07-960-932-8	Sequence 8, Appl
24	32.6	2.1	2817	1	US-07-960-932-9	Sequence 9, Appl
25	32.6	2.1	2861	1	US-07-960-932-10	Sequence 10, Appl
26	32.6	2.1	2794	1	US-07-908-253-1	Sequence 1, Appl
27	32.6	2.1	3311	1	US-08-170-126-1	Sequence 1, Appl
28	32.6	2.1	3329	1	US-08-170-126-3	Sequence 3, Appl
29	32.6	2.1	2794	2	US-08-455-970A-1	Sequence 1, Appl
30	32.6	2.1	2838	2	US-08-455-970A-9	Sequence 9, Appl
31	32.6	2.1	2817	2	US-08-455-970A-11	Sequence 11, Appl
32	32.6	2.1	2861	2	US-08-455-970A-13	Sequence 13, Appl
33	32.6	2.1	2794	2	US-08-387-156-5	Sequence 5, Appl
34	32.6	2.1	2934	2	US-08-387-156-7	Sequence 7, Appl
35	32.6	2.1	2794	3	US-08-694-865-5	Sequence 5, Appl
36	32.6	2.1	2934	3	US-08-694-865-7	Sequence 7, Appl
37	32.6	2.1	2653	3	US-08-589-711-1	Sequence 1, Appl

## ALIGNMENTS

RESULT 1  
US-08-480-604A-25  
; Sequence 25, Application US/08480604A  
; Patent No. 5736139  
; GENERAL INFORMATION:  
; APPLICANT: KINK, JOHN A.  
; APPLICANT: THALEY, BROCE S.  
; APPLICANT: PADHYE, NISHA V.  
; APPLICANT: FIRCA, JOSEPH R.  
; APPLICANT: STAFFORD, DOUGLAS C.  
; TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND  
; PREVENTION OF C. DIFFICILE DISEASE  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MEDLEN & CARROLL, LLP  
; STREET: 220 MONTGOMERY STREET, SUITE 2200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/480,604A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/422,711  
; FILING DATE: 14-APR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/405,496  
; FILING DATE: 16-MAR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/329,154  
; FILING DATE: 25-OCT-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/161,907  
; FILING DATE: 02-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/985,321  
; FILING DATE: 04-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/429,791  
; FILING DATE: 31-OCT-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: INGOLIA, DIANE E.  
; REGISTRATION NUMBER: 40,027  
; REFERENCE/DOCKET NUMBER: OPMD-01763  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1402 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear





```
LOCATION: 1..4362
PCT-US91-08525-31

Query Match
Best Local Similarity 47.6%; Pred. No. 0.27;
Matches 110; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 626 AGGATCTAGAAAAAGTTAGGCGGAGACACTTGTGCTCAGTTCTTTCAGTTTGAAG 685
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68 ATGATTGTAGACAAAGTTAAGGACAACTAGCTGCAATGAAACCTTATTAGAGACT 127
QY 686 AATTGCTCAGTTAGTCAAGATATAATATAGTATTTCCATTAAATATGATCCCGAA 745
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 128 TTTTGTTCAAAGTTTAAAGAGAGAAATGTAGTTGTTGGTGTATTATACCTCAG 187
QY 746 AAGATCGAGGTTTTCACCAATAGATAATTAATGATGATGCAATTTGCTCAAGAAA 805
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 188 AGGTGTGTACAACTGCTCTAGAACAGCACTACCACTGCCCTATGATATTTTAATA 247
QY 806 TCCTAGCTTATTTCTACCCGAGATACCATTTCTTAAAGCGGTCATTATG 856
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 248 TACATGCCCTTTTATTGATATGAGAGCTATGAAATAAGACTGTATATG 298

RESULT 6
PCT-US93-04384-11
; Sequence 11, Application PC/TUS9304384
; GENERAL INFORMATION:
; APPLICANT: Miller, Timothy J.
; APPLICANT: Klepfer, Sharon
; APPLICANT: Reed, Albert Paul
; APPLICANT: Jones, Elaine V.
; TITLE OF INVENTION: Compositions and Methods for Vaccination
; TITLE OF INVENTION: Against Coronaviruses
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithline Beecham Corporation - Corporate
; ADDRESSEE: Patents
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/04384
; FILING DATE: 19930507
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/882,171
; FILING DATE: 08-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/698,927
; FILING DATE: 13-MAY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/613,066
; FILING DATE: 14-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Schreck, Patricia A.
; REGISTRATION NUMBER: 33,777
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 270-5015
; TELEFAX: (215) 270-5090
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4365 base pairs
; TYPE: NUCLEIC ACID
```

```
STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4362
PCT-US93-04384-11

Query Match
Best Local Similarity 47.6%; Pred. No. 0.27;
Matches 110; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 626 AGGATCTAGAAAAAGTTAGGCGGAGACACTTGTGCTCAGTTCTTTCAGTTTGAAG 685
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68 ATGATTGTAGACAAAGTTAAGGACAACTAGCTGCAATGAAACCTTATTAGAGACT 127
QY 686 AATTGCTCAGTTAGTCAAGATATAATATAGTATTTCCATTAAATATGATCCCGAA 745
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 128 TTTTGTTCAAAGTTTAAAGAGAGAAATGTAGTTGTTGGTGTATTATACCTCAG 187
QY 746 AAGATCGAGGTTTTCACCAATAGATAATTAATGATGATGCAATTTGCTCAAGAAA 805
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 188 AGGTGTGTACAACTGCTCTAGAACAGCACTACCACTGCCCTATGATATTTTAATA 247
QY 806 TCCTAGCTTATTTCTACCCGAGATACCATTTCTTAAAGCGGTCATTATG 856
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 248 TACATGCCCTTTTATTGATATGAGAGCTATGAAATAAGACTGTATATG 298

RESULT 7
US-08-600-783-15
; Sequence 15, Application US/08600783
; Patent No. 5962267
; GENERAL INFORMATION:
; APPLICANT: SHIN, Hang Cheol
; APPLICANT: CHANG, Seung Gu
; APPLICANT: KIM, Dae Young
; APPLICANT: KIM, Chong Suh1
; TITLE OF INVENTION: Proinsulin Derivative and Process
; TITLE OF INVENTION: for Producing Human Insulin
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SHIN, Hang Cheol
; STREET: Ssangma-Hanshin Apt. 102-1206,
; STREET: #245 Cholsan-dong
; CITY: Kwangmyung-shi
; STATE: Kyungki-do
; COUNTRY: Republic of Korea
; ZIP: 423-030
; ADDRESSEE: CHANG, Seung Gu
; STREET: Hyundai Apt. 71-203, Apkujong-dong,
; STREET: Kangnam-ku
; CITY: Seoul
; STATE: Seoul
; COUNTRY: Republic of Korea
; ZIP: 135-110
; ADDRESSEE: KIM, Dae Young
; STREET: Sosa Jukong Apt. 108-202, Sosa Bon-dong,
; STREET: Sosa-ku
; CITY: Bucheon-shi
; STATE: Kyungki-do
; COUNTRY: Republic of Korea
; ZIP: 422-230
; ADDRESSEE: KIM, Chong Suh1
; STREET: Garden Heights Apt. 202-801, #100,
; STREET: Hwangkeum-dong, Soosung-ku
; CITY: Taegu
; STATE: Taegu
; COUNTRY: Republic of Korea
; ZIP: 706-040
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk, 3.5 inch, 1.44MB storage
```



Db 877 ACAATTATCATGATTAAGTTCCTGTCCTCCCAATAATTTATCTATTATCTATT 818  
QY 1195 GAAGAGATTTTAAAGCCAGCGAGACAGTACAAATTTCTCGAGAAAATGCCCTCAACGCC 1254  
Db 817 GTTTTTCATTTTAAGATTTGATTAATAACAAAACGACAAATAGTAATACATC 758  
QY 1255 ATTGAGTGATGGAGCGAGA 1276  
Db 757 ATGAACCTTAATGACCGAGATA 736

RESULT 10  
US-08-179-557-20  
Sequence 20, Application US/08179557  
Patent No. 5837509  
GENERAL INFORMATION:  
APPLICANT: ISRAELSEN, Hans  
APPLICANT: BECH HANSEN, Egon  
APPLICANT: MADSEN, Soeren Michael  
APPLICANT: JOHANSEN, Eric  
APPLICANT: NILSSON, Dan  
APPLICANT: VRANG, Astrid  
TITLE OF INVENTION: Recombinant Lactic Acid Bacterium  
TITLE OF INVENTION: Containing an inserted Promoter and Method of Constructing  
TITLE OF INVENTION: Same  
NUMBER OF SEQUENCES: 39  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W.  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/179,557  
FILING DATE: 07-JAN-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DK 1579/92  
FILING DATE: 30-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DK 0988/93  
FILING DATE: 01-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/DK94/00004  
FILING DATE: 03-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/036,681  
FILING DATE: 25-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30307/140/PLVI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 672 5300  
TELEFAX: 202 672 5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1100 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..57  
US-08-179-557-20

Query Match 2.2%; Score 34.6; DB 3; Length 1100;  
Best Local Similarity 48.7%; Pred. No. 0.87;  
Matches 94; Conservative 0; Mismatches 99; Indels 0; Gaps 0;  
QY 607 AACCCACAACATTTTATGAGATCTAGAAAAGTTAGGCTGGAACAACACTACTGTCAT 666  
Db 10 AAACACAGAGAAAATAAGAAAAATCAGAAAGATCAAAACACCCGAATAGGCTGTTTTCT 69  
QY 667 GGTTCCTCAGTTTGAAGAATTGGTTCAGTTAGTCAAGATAAATAATATAGATTTCC 726  
Db 70 TTTTATATGTTTATAGATAAGTGTCTAGTTATTTCTTGACAAAATAATATTTGAT 129  
QY 727 ATTAATATGATCCAGAAAAGATTGCGAGGTTTTCGCCAATAGATAATTAAGTATGAT 786  
Db 130 ATAAATTAATAGTTGTGCTTTGAGACGACGTACTTCTTATTAATTCATCAATAATATTAT 189  
QY 787 ATCGAATTGCCCA 799  
Db 190 TTGCAAAAGATTA 202

RESULT 11  
US-09-036-582-36  
Sequence 36, Application US/09036582A  
Patent No. 5965381  
GENERAL INFORMATION:  
APPLICANT: van der Bruggen, Pierre  
APPLICANT: Cornelis, Guy R.  
TITLE OF INVENTION: DELIVERY OF PROTEINS INTO EUKARYOTIC CELLS  
TITLE OF INVENTION: WITH RECOMBINANT YERSINIA  
FILE REFERENCE: 11154  
CURRENT APPLICATION NUMBER: US/09/036,582A  
CURRENT FILING DATE: 1998-03-06  
NUMBER OF SEQ ID NOS: 39  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 36  
LENGTH: 867  
TYPE: DNA  
ORGANISM: Yersinia enterocolitica  
US-09-036-582-36

Query Match 2.2%; Score 34; DB 4; Length 867;  
Best Local Similarity 44.3%; Pred. No. 1.1;  
Matches 139; Conservative 0; Mismatches 175; Indels 0; Gaps 0;  
QY 1010 CAATGATCATCATGATGATGCCCGTAGCAAGTTGCGCTGAAGATTAGCTGAGCTTACCG 1069  
Db 539 caaaaaaactttacaccgagagagatagcctgttgaataatacatgaagataataaaag 598  
QY 1070 CCGAATTAAGATTTATTCAGTTTTCAGCCGGAATTAATTAAGCATCTGCTAGTAGTG 1129  
Db 599 gtataatagtagtaggaataatcccttaccaccaataagttggatccgtatcccg 658  
QY 1130 GCACCATTAATTCATGATTAATTCATTAATTCATGATTAATTAATTAATTAATTAATTA 1189  
Db 659 taacttttacaacatcaccagaagtaaaaaagcttcaatgaatttaataataacc 718  
QY 1190 CAGATGAAGAGATTTTAAAGCCAGCGCAGAGTACAAAATTTTCGAGAAAATGCTTAAA 1249  
Db 719 cgcgggagttgtactgtttaaacaataaaatgaacaacatcttaataagtttgata 778  
QY 1250 CCACCATTCAGGTGATGGAGCGAGAAAATAATAGTCGATTAAGATTAAGATTTCTTGAA 1309  
Db 779 acaataatccattatagatgaaagaaatcattcaagtttcggtacacataaaagaaatag 838  
QY 1310 GTGAGATTAAGA 1323  
Db 839 ctgaataataaaca 852

RESULT 12

```

US-08-206-176-5
; Sequence 5, Application US/08206176
; Patent No. 5639940
; GENERAL INFORMATION:
; APPLICANT: Garner, Ian
; APPLICANT: Dalmtry, Michael A
; APPLICANT: Prunkard, Donna E
; APPLICANT: Foster, Donald C
; TITLE OF INVENTION: Production of fibrinogen in transgenic
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zymogenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/206.176
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 93-15
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10564 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; MOLECULE TYPE: linear
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: human fibrinogen gamma chain
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1799..1876, 1973..2017, 2207..2390, 2510
; LOCATION: ..2603, 4211..4341, 4645..4778, 5758..5942, 7426
; LOCATION: ..7703, 9342..9571)
US-08-206-176-5

Query Match          2.1%: Score 33.2; DB 1: Length 10564;
Best Local Similarity 51.3%: Pred. No. 7.3;
Matches 77: Conservative 0; Mismatches 73; Indels 0; Gaps 0;

OY 688 TTGGTTCAGTTGTCAGATTAATAATATAGATATTTCCATTAATAATATGATCCAGAAA 747
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3996 TTTGATTTGTTTCTTTTTCACAAAAGTTATATATTTTATTCATGTTATTTAGTAATA 4055
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

OY 748 GATTCGAGGTTTTCGCAATAGATTAATCTGATGATATGAAATGCTCAAGAAATC 807
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4056 TAAATTTACATTTTCCTCAAGATGAATATTTATTCAGAAAGCACTCTTAAGAAATA 4115
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

OY 808 CTAGCTTATTTCTACCCGAGATACCAT 837
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4116 CTTACGAGTTTCCAAAGAAATATTAAT 4145
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-08-323-170B-1
; Sequence 1, Application US/08323170B
; Patent No. 5733772
; GENERAL INFORMATION:

```

```

; APPLICANT: Williamson, Kim C.
; APPLICANT: Kaslow, David C.
; TITLE OF INVENTION: Cloning and Expression of Plasmidum
; TITLE OF INVENTION: fclilcparum Transmission-Blocking Target Antigen, Pfs230
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor.
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/323.170B
; FILING DATE: 13-OCT-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/010,409
; FILING DATE: 29-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Quine, Jonathan A.
; REGISTRATION NUMBER: P-41,261
; REFERENCE/DOCKET NUMBER: 015280-113100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 149..9556
US-08-323-170B-1

Query Match          2.1%: Score 33.2; DB 2: Length 9636;
Best Local Similarity 52.1%: Pred. No. 6.9;
Matches 74: Conservative 0; Mismatches 68; Indels 0; Gaps 0;

OY 283 TATAAAGAGGAGACACATCTGTTAACTTTACAGATGCGCGGGGATGCCATG 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7139 TTTAAATATCATTTATTTACAGCATATCTTAAGTACCGAGATCTTTAATTTATTT 7198
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

OY 343 TACTTAACATTTACTCTCAGAGATGAATACCAACCAATTCATCAACAAAGTATGCG 402
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7199 AACTCTCATGTAATTTGTTAATATCCAGAAAGCATGATATGATATGATATGCAATG 7258
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

OY 403 AAGATCTCAGAGATTTGATTA 424
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7259 TCTGATTAATCGAAATTTGATA 7280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
US-08-229-145-13
; Sequence 13, Application US/08229145
; Patent No. 5691461
; GENERAL INFORMATION:
; APPLICANT: Hoke, Glenn D.
; APPLICANT: Ecker, David J.
; TITLE OF INVENTION: INHIBITION OF CANDIDA USING
; TITLE OF INVENTION: OLIGONUCLEOTIDES
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Macklewitz &

```







```
221 GHGSSVLEELVOLVKDKNIDISIKIDPRKDSYFANRVITDIELLKITL 270
|||||
205 GHGSSVLEELVOLVKDKNIDISIKIDPRKDSYFANRVITDIELLKITL 254
|||||
271 AFLEPDLITLKGHYDNOLONGIKRYKEFLLESSPNTQWELRAFMAVMHFS 320
|||||
255 AFLEPDLITLKGHYDNOLONGIKRYKEFLLESSPNTQWELRAFMAVMHFS 304
|||||
321 LFADRIDDDILKVIYDSNMNHGDKRSKLRELAELTAELKIVSVIOAEIN 370
|||||
305 LFADRIDDDILKVIYDSNMNHGDKRSKLRELAELTAELKIVSVIOAEIN 354
|||||
371 KHLSSGTTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIQ 420
|||||
355 KHLSSGTTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIQ 404
|||||
421 VGSSEKKIYISIKDFLGSSEKRTGALGNLKNYSYNKDNNELSHFATTCSD 470
|||||
405 VGSSEKKIYISIKDFLGSSEKRTGALGNLKNYSYNKDNNELSHFATTCSD 454
|||||
471 KSRPLNDLVSOQKTOLSDITSRFSNAIEALNRFIOKYDSVMORLLDDTSG 520
|||||
455 KSRPLNDLVSOQKTOLSDITSRFSNAIEALNRFIOKYDSVMORLLDDTSG 504
|||||
521 K 521
1
505 K 505

seq_name: gb_pat:A56808

seq_documentation_block:
LOCUS A56808 1462 bp DNA PAT 03-MAR-1998
DEFINITION Sequence 16 from Patent WO9628551.
ACCESSION A56808
NID g3712821
VERSION A56808.1 GI:3712821
KEYWORDS
SOURCE Yersinia pestis.
ORGANISM Yersinia pestis
Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;
Yersinia.
REFERENCE 1 (bases 1 to 1462)
AUTHORS Tlball,R.W., Williamson,E.D., Leary, Sophie,E.C., Oyston,P.C.,
Bennett and Alice,M.
TITLE VACCINES FOR PLAGUE
JOURNAL Patent: WO 9628551-A 16 19-SEP-1996;
SECR DEFENCE (GB)
COMMENT Other publication ZA 9602036 960716
FEATURES
source location AU 4951196 961002.
1..1462
/organism="Yersinia pestis"
/db_xref="taxon:632"
CDS
<8..1450
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CA034321.1"
/db_xref="PID:e1255331"
/db_xref="PID:g3712822"
/translation="ADLTASTTATATLVEPARITLTKEGAPITIMDNGNIDELVLG
TLTLGGYKGTSTSVNFTDAAGDPMTLFTSODGNHOFRTTVIGKDSRDPISPY
NGENLVGDDVVLATGSDQDFVRSIGSGKGLAAGKYDAVTYVNSGSIKGRIRAE
ONRPHFIEDEKRVEDLTGHGSSVLEELVOLVKDKNIDISIKIDPRKDSYFANRV
ITDIELLKITLAFLEPDLITLKGHYDNOLONGIKRYKEFLLESSPNTQWELRAFMAV
HFS/LFADRIDDDILKVIYDSNMNHGDKRSKLRELAELTAELKIVSVIOAEINHLSS
SGTTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIOVDSSEKKIYSIKD
FLGSSEKRTGALGNLKNYSYNKDNNELSHFATTCSDKSRPLNDLVSOQKTOLSDITS
RFSNAIEALNRFIOKYDSVMORLLDDTSG"
```

```
alignment_scores:
  Quality: 2033.00 Length: 482
  Ratio: 4.262 Gaps: 1
  Percent Similarity: 98.963 Percent Identity: 98.340
```

```
alignment_block:
us-08-699-716a-2 x A56808 (2) ..
```

```
Align seg 1/1 to: A56808 from: 1 to: 1462
```

```
43 NADLTASTTATATLVEPARITLTKEGAPITIMDNGNIDELVLGTLT 92
:::
1 SSADLTASTTATATLVEPARITLTKEGAPITIMDNGNIDELVLGTLT 50
93 GGYKGTSTSVNFTDAAGDPMTLFTSODGNHOFRTTVIGKDSRDPDI 142
|||||
51 GGYKGTSTSVNFTDAAGDPMTLFTSODGNHOFRTTVIGKDSRDPDI 100
|||||
143 SPKVENGLVGDVVLATGSDQDFVRSIGSGKGLAAGKYDAVTYVSN 192
|||||
101 SPKVENGLVGDVVLATGSDQDFVRSIGSGKGLAAGKYDAVTYVSN 150
|||||
193 Q...EFMIRAYEQNPQHFIEDEKRVEDLTGHGSSVLEELVOLVKDKNI 239
1
151 QGSIKGRIRAYEQNPQHFIEDEKRVEDLTGHGSSVLEELVOLVKDKNI 200
|||||
240 DISIKYDPRKDSYFANRVITDIELLKITLAFLEPDLITLKGHYDNOL 289
|||||
201 DISIKYDPRKDSYFANRVITDIELLKITLAFLEPDLITLKGHYDNOL 250
|||||
290 QNGIKRYKEFLLESSPNTQWELRAFMAVMHFS/LFADRIDDDILKVIYDSMN 339
|||||
251 QNGIKRYKEFLLESSPNTQWELRAFMAVMHFS/LFADRIDDDILKVIYDSMN 300
|||||
340 HHGDARSKLRELAELTAELKIVSVIOAEINKHLSSGTTINIHDKSINLM 389
|||||
301 HHGDARSKLRELAELTAELKIVSVIOAEINKHLSSGTTINIHDKSINLM 350
|||||
390 DNKLGYTDEEIFKASAEYKILEKMPOTTIOVDSSEKKIYISIKDFLGSSE 439
|||||
351 DNKLGYTDEEIFKASAEYKILEKMPOTTIOVDSSEKKIYISIKDFLGSSE 400
|||||
440 KRTGALGNLKNYSYNKDNNELSHFATTCSDKSRPLNDLVSOQKTOLSDI 489
|||||
401 KRTGALGNLKNYSYNKDNNELSHFATTCSDKSRPLNDLVSOQKTOLSDI 450
|||||
490 TSRFNSAIEALNRFIOKYDSVMORLLDDTSGK 521
|||||
451 TSRFNSAIEALNRFIOKYDSVMORLLDDTSGK 482

seq_name: gb_bal:YEPICR

seq_documentation_block:
LOCUS YEPICR 2100 bp DNA BCT 26-APR-1993
DEFINITION Yersinia pestis lcrG, lcrV, and lcrH genes, complete cds.
ACCESSION M26405
NID g155448
VERSION M26405.1 GI:155448
KEYWORDS lcrG protein; lcrH protein; lcrV protein; V antigen.
SOURCE 75kb virulence plasmid.
ORGANISM Yersinia pestis
Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;
Yersinia.
REFERENCE 1 (bases 1 to 2100)
AUTHORS Price,S.B., Leung,K.Y., Barve,S.S. and Straley,S.C.
TITLE Molecular analysis of lcrGVH, the V antigen operon of Yersinia
JOURNAL J. Bacteriol. 171, 5646-5653 (1989)
MEDLINE 90008806
COMMENT Computer readable copy of sequence [J. Bacteriol. (1989) In press]
kindly submitted by Price,S.
```

```
FEATURES
    07-AUG-1989.
Source      Location/Qualifiers
            1..2100
            /organism="Yersinia pestis"
            /strain="KIM5"
            /db_xref="taxon:632"
-35_signal
-10_signal
gene
    147..152
    174..179
    192..479
    /gene="lcrG"
    192..479
    /gene="lcrG"
    /codon_start=1
    /transl_table=11
    /protein_id="AA27640.1"
    /db_xref="GI:155449"
    /db_xref="PID:g155449"
    /translation="MKSHFDEYDKTLKQAEIAIADSHRAKILQMCADIGLTPEAV
    MKIFAGRSAEIKPAERELDEIKROREROPHPYDGKRKPTMRGOI"
    463..466
    /gene="lcrG"
    463..1461
    /gene="lcrV"
    481..1461
    /gene="lcrV"
    /codon_start=1
    /transl_table=11
    /protein_id="AA27641.1"
    /db_xref="PID:g155450"
    /db_xref="GI:155450"
    /translation="MIRAYEONPOHFIEDLEKRVBOLTGHGSSVLEELVOLVKDNI
    DISIKYDPKRDSEVFANRVTDDIELLKILAYFLPEDAILKGHYDNOLONGIKRVK
    EFLSESPNTOEWELRAFMAVHFSLTADRIDDDILKVIYDSNMHNGDARSKLREEL
    TELKIVSYIOAEINKHLSGGTINIHDKSINLMDKNLYG
    MROTTIÖVDSSEKKIYSIKDFLGSSEKRRKGALGNLKNSTSYKDNNELSHFATCSDK
    SRPLNDVSOQKTTÖLSDITSRENSAIEALNRFÖKYDSVÖRLDDTSGK"
    1465..1468
    /gene="lcrH"
    1465..1980
    /gene="lcrH"
    1474..1980
    /gene="lcrH"
    /codon_start=1
    /transl_table=11
    /protein_id="AA27642.1"
    /db_xref="PID:g155451"
    /db_xref="GI:155451"
    /translation="MOEFTTDOEYOLAMESFLKGGGTTAMLNEISSDTLEOLYSLAF
    NQYOSKYEDAHRVFOALCVLDHYDSRPFILGACROAMGOYDLAIHSYSYGIMDIK
    EPRFPFHAECCLÖQKGEIAEESGLFLAÖELLADKTEPFELSTRVSSMLEAIKTKREM
    EHECVDPN"
BASE COUNT      688 a      423 c      461 g      528 t
ORIGIN
alignment_scores:
    Quality: 1450.00      Length: 326
    Ratio: 4.448          Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 99.693
alignment_block:
US-08-699-716A-2 x YEPLCR (1)
Align seg 1/1 to: YEPLCR from: 1 to: 2100
196 MIRAYEONPOHFIEDLEKRVBOLTGHGSSVLEELVOLVKKNIDISIKY 245
|||||
161 MIRAYEONPOHFIEDLEKRVBOLTGHGSSVLEELVOLVKKNIDISIKY 210
|||||
246 DPRKDSVFANRVTDDIELLKILAYFLPEDTILKGHYDNOLONGIKR 295
|||||
211 DPRKDSVFANRVTDDIELLKILAYFLPEDAILKGHYDNOLONGIKR 260
|||||
```

```
296 VKFELESSPNTÖWELRAFMAVHFSLTADRIDDDILKVIYDSNMHGDAR 345
|||||
261 VKFELESSPNTÖWELRAFMAVHFSLTADRIDDDILKVIYDSNMHGDAR 310
|||||
346 SKLRBEELAEELKILYSVIOAEINKHLSGGTINIHDKSINLMDKNLYG 395
|||||
311 SKLRBEELAEELKILYSVIOAEINKHLSGGTINIHDKSINLMDKNLYG 360
|||||
396 YTDEIFRASAEEYKILEKMPÖTTIÖVDSSEKKIYSIKDFLGSSEKRRKGAL 445
|||||
361 YTDEIFRASAEEYKILEKMPÖTTIÖVDSSEKKIYSIKDFLGSSEKRRKGAL 410
|||||
446 GNLSNSYSYKDNNELSHFATCSDKSRPLNDVSOQKTTÖLSDITSRFSNS 495
|||||
411 GNLSNSYSYKDNNELSHFATCSDKSRPLNDVSOQKTTÖLSDITSRFSNS 460
|||||
496 AIEALNRFÖKYDSVÖRLDDTSGK 521
|||||
461 AIEALNRFÖKYDSVÖRLDDTSGK 486
|||||
seq_name: gb_pat:A46411
seq_documentation_block:
LOCUS      A46411      1014 bp      DNA
DEFINITION Sequence 1 from Patent WO9524475.
ACCESSION      A46411
NID      92300612
VERSION      A46411.1 GI:2300612
KEYWORDS
SOURCE
ORGANISM      Yersinia pestis.
              Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
              Yersinia pestis
REFERENCE
1 (bases 1 to 1014)
AUTHORS      Tiddall,R.W., Williamson,E.D. and Leary,S.E.
TITLE      VACCINE COMPOSITIONS
JOURNAL      Patent: WO 9524475-A 1 14-SEP-1995;
              SECR DEFENCE BRIT (GB)
COMMENT      Other publication AU 1853995 950925.
FEATURES
    source
        1..1014
        /organism="Yersinia pestis"
        /db_xref="taxon:632"
        1..990
        /note="unnamed protein product"
        /codon_start=1
        /transl_table=11
        /protein_id="CA002893.1"
        /db_xref="PID:g306155"
        /db_xref="PID:g2300613"
        /db_xref="GI:2300613"
        /translation="MSEIRAYEONPOHFIEDLEKRVBOLTGHGSSVLEELVOLVKD
        KNIDISIKYDPKRDSEVFANRVTDDIELLKILAYFLPEDAILKGHYDNOLONGIK
        RVKEFLSESPNTOEWELRAFMAVHFSLTADRIDDDILKVIYDSNMHNGDARSKLREEL
        AELTELKIVSYIOAEINKHLSGGTINIHDKSINLMDKNLYGTYDEIFRASAEEYKI
        LEKMPÖTTIÖVDSSEKKIYSIKDFLGSSEKRRKGALGNLKNSTSYKDNNELSHFATTC
        SRSRPLNDVSOQKTTÖLSDITSRENSAIEALNRFÖKYDSVÖRLDDTSGK"
BASE COUNT      346 a      181 c      201 g      286 t
ORIGIN
alignment_scores:
    Quality: 1443.00      Length: 325
    Ratio: 4.440          Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 99.692
alignment_block:
US-08-699-716A-2 x A46411 (1)
Align seg 1/1 to: A46411 from: 1 to: 1014
197 IIRAYEONPOHFIEDLEKRVBOLTGHGSSVLEELVOLVKKNIDISIKYD 246
|||||
```

```
5 IRAYEONPOHFIEDLEKRVVEOLTGHSSVLEELVOLVKDKNIDISIKYD 54
247 PRKDEVFANRVITDDIELKKILAYFLPEPTILKGHYDNOLONGIKRV 296
55 PRKDSVFANRVITDDIELKKILAYFLPEDAILKGHDNOLONGIKRV 104
297 KEFLSSPNTOWELRAFMAVMHFSLTADRIDDILKVIYDSNMHHGDARS 346
105 KEFLSSPNTOWELRAFMAVMHFSLTADRIDDILKVIYDSNMHHGDARS 154
347 KREBELAELTAEIKIYSVIOAEINKHLSSSGTINIHDKSINLMDKNLYGY 396
155 KREBELAELTAEIKIYSVIOAEINKHLSSSGTINIHDKSINLMDKNLYGY 204
397 TDEEIFKASAEYKILEKMPOTTIOVGSEKKIVSIKDFLGSENKRTGALG 446
205 TDEEIFKASAEYKILEKMPOTTIOVGSEKKIVSIKDFLGSENKRTGALG 254
447 NLKNSYSYKNDNNELSHFATCSDKSRPLNDLVSOQTQLSDITSRFNSA 496
255 NLKNSYSYKNDNNELSHFATCSDKSRPLNDLVSOQTQLSDITSRFNSA 304
497 IEALNRFIOKYDSVMQRLDDTSGK 521
305 IEALNRFIOKYDSVMQRLDDTSGK 329
```

seq\_name: gb\_pat:A46413

seq\_documentation\_block:

LOCUS A46413 1014 bp DNA PAT 07-MAR-1997  
DEFINITION Sequence 3 from Patent WO9524475.  
ACCESSION A46413  
NID 92300614  
VERSION A46413.1 GI:2300614

KEYWORDS  
SOURCE

ORGANISM

REFERENCE

AUTHORS 1 (bases 1 to 1014)  
TITLE Tiltball,R.W., Williamson,E.D. and Leary,S.E.  
JOURNAL VACCINE COMPOSITIONS  
Patent: WO 9524475-A 3 14-SEP-1995;  
SECR DEFENCE BRIT (GB)  
Other publication AU 1853995 950925.  
Location/Qualifiers

COMMENT  
FEATURES

source

CDS

1. 1014  
/organism="Yersinia pestis"  
/db\_xref="taxon:632"  
1. 987  
/note="unnamed protein product; protein sequence is in  
conflict with the conceptual translation"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CA02894.1"  
/db\_xref="PID:e306384"  
/db\_xref="PID:g2300615"  
/db\_xref="GI:2300615"  
/translation="GIRIRAYEONPOHFIEDLEKRVVEOLTGHSSVLEELVOLVKD  
KNIDISIKYDPRKDSVFANRVITDDIELKKILAYFLPEDAILKGHDNOLONGIK  
RVEFLSSPNTOWELRAFMAVMHFSLTADRIDDILKVIYDSNMHHGDARSKIREL  
AELTAEIKIYSVIOAEINKHLSSSGTINIHDKSINLMDKNLYGYTDEEIFKASAEYK  
LEKMPOTTIOVGSEKKIVSIKDFLGSENKRTGALGINKNSYSYKNDNNELSHFATC  
SDKSRPLNDLVSOQTQLSDITSRFNSAIEALNRFIOKYDSVMQRLDDTSGK"  
BASE COUNT 343 a 184 c 205 g 282 t

alignment\_scores:

Quality: 1443.00 Length: 325  
Ratio: 4.440 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 99.692

alignment\_block:  
US-08-699-716a-2 x A46413 (1)  
Align seg 1/1 to: A46413 from: 1 to: 1014

```
197 IRAYEONPOHFIEDLEKRVVEOLTGHSSVLEELVOLVKDKNIDISIKYD 246
|||||
5 IRAYEONPOHFIEDLEKRVVEOLTGHSSVLEELVOLVKDKNIDISIKYD 54
247 PRKDEVFANRVITDDIELKKILAYFLPEPTILKGHYDNOLONGIKRV 296
55 PRKDSVFANRVITDDIELKKILAYFLPEDAILKGHDNOLONGIKRV 104
297 KEFLSSPNTOWELRAFMAVMHFSLTADRIDDILKVIYDSNMHHGDARS 346
105 KEFLSSPNTOWELRAFMAVMHFSLTADRIDDILKVIYDSNMHHGDARS 154
347 KREBELAELTAEIKIYSVIOAEINKHLSSSGTINIHDKSINLMDKNLYGY 396
155 KREBELAELTAEIKIYSVIOAEINKHLSSSGTINIHDKSINLMDKNLYGY 204
397 TDEEIFKASAEYKILEKMPOTTIOVGSEKKIVSIKDFLGSENKRTGALG 446
205 TDEEIFKASAEYKILEKMPOTTIOVGSEKKIVSIKDFLGSENKRTGALG 254
447 NLKNSYSYKNDNNELSHFATCSDKSRPLNDLVSOQTQLSDITSRFNSA 496
255 NLKNSYSYKNDNNELSHFATCSDKSRPLNDLVSOQTQLSDITSRFNSA 304
497 IEALNRFIOKYDSVMQRLDDTSGK 521
305 IEALNRFIOKYDSVMQRLDDTSGK 329
```

seq\_name: gb\_pat:A56793

seq\_documentation\_block:

LOCUS A56793 1014 bp DNA PAT 03-MAR-1998  
DEFINITION Sequence 1 from Patent WO9628551.  
ACCESSION A56793  
NID 93712808  
VERSION A56793.1 GI:3712808

KEYWORDS  
SOURCE

ORGANISM

REFERENCE 1 (bases 1 to 1014)  
AUTHORS Tiltball,R.W., Williamson,E.D., Leary, Sophie,E.C., Oyston,P.C.,  
Bennett and Alice,M.  
TITLE VACCINES FOR PLAGUE  
JOURNAL Patent: WO 9628551-A 1 19-SEP-1996;  
SECR DEFENCE (GB)  
Other publication ZA 9602036 960716  
Other publication AU 4951196 961002.  
Location/Qualifiers

COMMENT  
FEATURES

source

CDS

1. 1014  
/organism="Yersinia pestis"  
/db\_xref="taxon:632"  
1. 990  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CA03419.1"  
/db\_xref="PID:e125515"  
/db\_xref="PID:g4530021"  
/db\_xref="GI:4530021"  
/translation="ISEFIRAYEONPOHFIEDLEKRVVEOLTGHSSVLEELVOLVKD  
KNIDISIKYDPRKDSVFANRVITDDIELKKILAYFLPEDAILKGHDNOLONGIK  
RVEFLSSPNTOWELRAFMAVMHFSLTADRIDDILKVIYDSNMHHGDARSKIREL  
AELTAEIKIYSVIOAEINKHLSSSGTINIHDKSINLMDKNLYGYTDEEIFKASAEYK  
LEKMPOTTIOVGSEKKIVSIKDFLGSENKRTGALGINKNSYSYKNDNNELSHFATC  
SDKSRPLNDLVSOQTQLSDITSRFNSAIEALNRFIOKYDSVMQRLDDTSGK"  
BASE COUNT 346 a 181 c 201 g 286 t



JOURNAL

Submitted (03-SEP-1999) Yersinia pestis sequencing project, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: barrell@sanger.ac.uk DNA supplied by Dr. Andrey Karlyshev and Prof. Brendan Wren, [3]. Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT

COMMENT

Notes: Yersinia pestis sequencing at The Sanger Centre is funded by

Beowulf Genomics Details of Y. pestis sequencing at the Sanger Centre are available on the World Wide Web.

(URL: <http://www.sanger.ac.uk/Projects/Y-pestis/>) CDS are numbered using the following system eg YPCP1.01c, YP (Y. pestis), PCP1 (plasmid name), .01 (first CDS), c (complementary strand).

The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous. The length in codons is given for each CDS.

Usually the highest scoring match found by fasta\_0 is given for

CDS which show significant similarity to other CDS in the database. The position of possible ribosome binding site sequences are given

where these have been used to deduce the initiation codon. CAUTION: We may not have predicted the correct initiation codon.

Where possible we choose an initiation codon (atg, gtg, ttg or att) which is preceded by an upstream ribosome binding site

sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

FEATURES

source

Location/Qualifiers

1..70305

/organism="Yersinia pestis"

/plasmid="PCD1"

/strain="CO-92 Biovar Orientalis"

/db\_xref="taxon:632"

1..1956

/note="IS100 element"

88..1110

/gene="YPCD1.01"

88..1110

/gene="YPCD1.01"

88..1110

/gene="YPCD1.01"

88..1110

/gene="YPCD1.01"

88..1110

/gene="YPCD1.01"

88..1110

/gene="YPCD1.01"

88..1110

/gene="YPCD1.01"

88..1110

/gene="YPCD1.01"

88..1110

/gene="YPCD1.01"

88..1110

/gene="YPCD1.01"

88..1110

/gene="YPCD1.01"

88..1110

/gene="YPCD1.01"

88..1110

/gene="YPCD1.01"

88..1110

/gene="YPCD1.01"

88..1110

/gene="YPCD1.01"

88..1110

/gene="YPCD1.01"

88..1110

/gene="YPCD1.01"

88..1110

/gene="YPCD1.01"

88..1110

/gene="YPCD1.01"

88..1110

/gene="YPCD1.01"

88..1110

misc\_feature 2304..2982  
/note="similar to E.agglomerans IS 1222 (EMBL:X78052) at DNA level"  
gene complement(3014..3406)  
/gene="YPCD1.05c"  
CDS complement(3014..3406)  
/gene="YPCD1.05c"  
/note="YPCD1.05c, syce, yerA, yope targeting protein, len: 130 aa, identical to corresponding CDS from Y.pestis KIM5 pCD1 (EMBL:AF053946, AF074612) (130 aa), fasta scores: opt: 861 z-score: 1608.5 E(): 0, 100.0% identity in 130 aa overlap and to YERA.YERPE (EMBL:M34279) from Y.pestis plasmid pYV019. Highly similar to TR-Q56910 (EMBL:Z18539), syce, from Yersinia enterocolitica (130 aa) (99.28% identity in 130 aa overlap) and YERA.YEREN (EMBL:M34278), yerE, also from Y.enterocolitica (130 aa) (98.5% identity in 130 aa overlap). Similar to TR-Q51448 (EMBL:L27629), ORF1, Pseudomonas aeruginosa exoenzyme S ORF1 (116 aa) (44.3% identity in 115 aa overlap)"  
/codon\_start=1  
/transl\_table=11  
/label=syce  
/product="putative yope chaperone"  
/protein\_id="CA854882.1"  
/db\_xref="PID:ca854882.1"  
/db\_xref="GI:5832428"  
/translation="MYSEFOAITQLFOQLSLSDPTIEPIYGVKGEFACHTEHPYQ OIMFTPLSDNNDEKFTLISHNFSQDILKPLISNDEVGCHYVLMNRQPLNLDNNS LYTOLEMGAGARLOTSSLSPRRSS"  
gene 3600..4259  
/gene="YPCD1.06"  
CDS 3600..4259  
alignment\_scores:  
Quality: 1450.00 Length: 326  
Ratio: 4.448 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 99.693  
alignment\_block:  
US-08-699-716a-2 x YPCD1/rev (6) ..  
Align seg 1/1 to reverse of: YPCD1 from: 1 to: 70305  
196 MIRAEQNPNHFIEDLEKRVQLTGHGSSVLEELVQLYKKNIDISIKY 245  
|||||  
7638 MIRAEQNPNHFIEDLEKRVQLTGHGSSVLEELVQLYKKNIDISIKY 7589  
246 DPRKSEVFNANVTDDIELKKILAYFLPEDTILKGGHYDNOQNGIKR 295  
|||||  
7588 DPRKSEVFNANVTDDIELKKILAYFLPEDTILKGGHYDNOQNGIKR 7539  
296 VKPELESSPNQWELRAFVAVHFSLTADRIDDDILKIVDSMNHGDKR 345  
|||||  
7538 VKPELESSPNQWELRAFVAVHFSLTADRIDDDILKIVDSMNHGDKR 7489  
346 SKLRELAELTAELKIYSYIOAEINKHLSSTGINIHDKSIMLMDKNLYG 395  
|||||  
7488 SKLRELAELTAELKIYSYIOAEINKHLSSTGINIHDKSIMLMDKNLYG 7439  
396 YTDEEIFKASAEYKILEKMPOTTIOVDSSEKRVISIKDFLGSSENRRTGAL 445  
|||||  
7438 YTDEEIFKASAEYKILEKMPOTTIOVDSSEKRVISIKDFLGSSENRRTGAL 7389  
446 GULKKSYSYKNNLNSHFATCSDKSRPLNDLVQSKTQLDLSIRFNS 495  
|||||  
7388 GULKKSYSYKNNLNSHFATCSDKSRPLNDLVQSKTQLDLSIRFNS 7339  
496 AIEALNRFIQKYDSVMORLDDTSCK 521  
|||||  
7338 AIEALNRFIQKYDSVMORLDDTSCK 7313  
seq\_name: gb\_ba2:AF053946

seq\_documentation\_block:  
LOCUS AF053946 70504 bp DNA circular BCT 06-OCT-1998  
DEFINITION Yersinia pestis plasmid pCD1, complete plasmid sequence.  
ACCESSION AF053946  
NID 92996222  
VERSION AF053946.1 GI:2996222  
KEYWORDS Yersinia pestis.  
SOURCE Yersinia pestis  
ORGANISM Bacteria: Proteobacteria: gamma subdivision; Enterobacteriaceae; Yersinia.  
REFERENCE 1 (bases 1 to 70504)  
AUTHORS Hu,P., Elliott,J., McCready,P., Skowronski,E., Garnes,J., Kobayashi,A., Brubaker,R.R. and Garcia,E.  
TITLE Structural organization of virulence-associated plasmids of Yersinia pestis  
JOURNAL J Bacteriol. 180 (19), 5192-5202 (1998)  
MEDLINE 96422474  
REFERENCE 2 (bases 1 to 70504)  
AUTHORS Hu,P., Elliott,J., McCready,P., Skowronski,E., Garnes,J., Kobayashi,A., Carrano,A.V., Brubaker,R. and Garcia,E.  
TITLE Direct Submission  
JOURNAL Submitted (16-MAR-1998) Biology and Biotechnology, Lawrence Livermore National Lab, L452, 7000 East Ave., Livermore, CA 94550, USA  
FEATURES  
source Location/Qualifiers  
1..70504  
/organism="Yersinia pestis"  
/plasmid="pCD1"  
/strain="Kim"  
/db\_xref="taxon:632"  
87..1109  
/codon\_start=1  
/transl\_table=11  
/product="transposase"  
/protein\_id="AAC62546.1"  
/db\_xref="PID:g2996223"  
/db\_xref="GI:2996223"  
/translation="MYTPREAVASLIDETRDYIRORIANAHYKIPATYIAIEIRDOGITIRATIR SLSVPOEPAVAFETEPGRQOVQDNGRNGSPHVAIVAGYSRMUYIETDMR YDLTETCHRNAEFEEFGVPREVLYDNMKTIVLQDCAVOTGQHFHPSLMQFGKMGFS PLCRPFRATOKKRVEMVOYTERNSFYIPLMTRLRPMGITVDVETANRGLRLHDVA NORHETIQARPCDRLEEQSMALPPEKKEVDVLDENLVNFDKHLPHPLSTIDS FCGXVA"  
1109..1888  
/codon\_start=1  
/transl\_table=11  
/product="transposase"  
/protein\_id="AAC62557.1"  
/db\_xref="PID:g2996234"  
/db\_xref="GI:2996234"  
/translation="MMELQHORLMLALQLOLESLISAPALSOQAVDQWDSYMDLE HLHEEKRLAHQKQAMYTRMAAFPAVKFEEDYDFPAGAPQKQOSLRSLFIEEN ENVLVIGPSGVGKTHLAIAMGYAVRAGIKRRTTADLLIOSTNROGRKYTTIOR GVAAPRLIITDEIGYIPFQOEAKLFFOYIARKEYSAMITITNLFQGDQDTFAGDA ALTSAMDRLHSHSHVQIRGESYRURORRKAQVIAEAND"  
complement(1939..2343)  
/codon\_start=1  
/transl\_table=11  
/product="transposase"  
/protein\_id="AAC62568.1"  
/db\_xref="PID:g2996245"  
/db\_xref="GI:2996245"  
/translation="MTEYQASERRCRIMGISRSLHAYCPNTRDIPVVEVLOKLAHQ YPAVGGFMENKLRQGLPMNVKRVYRLKLTNRRGRKKRLPRRHQPLAIPLMK NHCWSDVMSDALTDGRRFRLENVVEIIFVNDG"  
complement(2379..2645)  
/gene="lcrS"  
/complement(2379..2645)  
/gene="lcrS"  
/codon\_start=1  
CDS

gene  
CDS .  
gene  
CDS  
gene  
CDS  
gene  
CDS  
gene  
CDS  
gene  
CDS

/transl\_table=11  
/product="low calcium response locus protein S homolog"  
/protein\_id="AAC62579.1"  
/db\_xref="PID:g2996256"  
/db\_xref="GI:2996256"  
/translation="MKKARFETQILPVLKEVEGRHVKDYCRENGCSSEASYNWKSK  
YGMSSSIIRKKREERNRRLKQVAYSLSDHELKLKVAKKI"  
complement(3193..3540)  
/gene="yscm"  
complement(3193..3540)  
/gene="yscm"  
/codon\_start=1  
/transl\_table=11  
/product="top proteins translocation protein M homolog"  
/protein\_id="AAC62586.1"  
/db\_xref="PID:g2996263"  
/db\_xref="GI:2996263"  
/translation="MKINTLOSLINOOITVOYGHGGOAGRLPETNPLENSHQISTAEK  
AFANVIEHVKNRTALSRIIDACLPRLPSNLBKKGKAGEIVITGLRTEQLSLSAKLIL  
LEAMRKDTADG"  
complement(3765..4430)  
/gene="yscl"  
complement(3765..4430)  
/gene="yscl"  
/codon\_start=1  
/transl\_table=11  
/product="top proteins translocation protein L homolog"  
/protein\_id="AAC62596.1"  
/db\_xref="PID:g2996273"  
/db\_xref="GI:2996273"  
/translation="MSOTCQTGYAMQPFVQIITSNSLACGLILKAEDYOSLTTE  
ELISAKQDAEKILADAQEVYEQQOLQGOMGEARTLQTLIHETLOQCQYRYHV  
EOKMEEVDLAARKILINDYQVAAMLGVRRALALYSNQKVVVRVPDGAJREDID  
AKVHFDFEISLYLEVTDARLDQGGCILETEYGIIDSIGDIALSRAISTTIGQKK  
VTE"  
complement(4376..5005)  
/gene="ysck"  
complement(4376..5005)  
/gene="ysck"  
/codon\_start=1  
/transl\_table=11  
/product="top proteins translocation protein K homolog"  
/protein\_id="AAC62605.1"  
/db\_xref="PID:g2996282"  
/db\_xref="GI:2996282"  
/translation="MMENVITSFOLFRCPPAAYLVHLDELPSPMSRIPLPYLPOMRSAILH  
NAALDEFSLDDVEEPHGICALPIPOPOSLELLCRGIVLHGFAIRCVLASPLDOO  
LLTLVNQETLRILYQHELLIGMPHTMQRELPIEFBSERTMIQSGLFWLAAAMPQFO  
AMCKLSLRPLATISEPWLVAESORPLAQTLCHRLKYQVPTCSHLK"  
complement(5005..5739)  
/gene="yscu"  
complement(5005..5739)  
/gene="yscu"  
complement(5005..5739)  
/gene="yscu"  
/codon\_start=1  
/transl\_table=11  
/product="top proteins translocation lipoprotein J"  
/protein\_id="AAC62607.1"  
/db\_xref="PID:g2996284"  
/db\_xref="GI:2996284"  
/translation="MKVKTSLSTLILILFLTGCKVYDTGTSGKEGNEMALLROBEL  
SADKRPDDGKITKLVEESDVQAQIDILKRKGYPHESTLQDPFPDGILTSPIEBEL  
ARLAPKDEISRLETSEDGVAVRHVAVLVLEDENNGKRGVASASVFIKHADIDOFD  
DYIIOIOLVANSIEGLADVRIISYLVPVSVDVROSSLPANTSILSIQVSEEEKHLH  
IGLSLILILLPVNTLAQYFMIQRKK"  
complement(5746..6093)  
/gene="ysci"  
complement(5746..6093)  
/gene="ysci"  
/codon\_start=1  
/transl\_table=11  
/product="top proteins translocation protein I homolog"  
/protein\_id="AAC62608.1"  
/db\_xref="PID:g2996285"

```

/gb_xref="GI:2996285"
/translation="MPNIEBAQADDEVITITTEELGPAEPTDQIMRFDAKSPDTQGL
GSHLKEVSDIKSKFTKVSQDLTKLAVSDNPNDMLMQMSLIRITIOELLAKTAG
RMSQNVETLSKSG"
complement(6094. .6591)
/gene="ysch"
/complement(6094. .6591)
/gene="ysch"
/codon_start=1
/transl_table=11
/product="yop proteins translocation protein H homolog"
/protein_id="AAC62547.1"
/db_xref="PID:g2996224"
/db_xref="GI:2996224"
/translation="MTVTLNRGSIITSLMSSQAVSTLQPVASELTKQLENKLSSEAE
KTRFVLMQOYYASNPDDHAVLEVLATPVEPALARGOHSVPAIDIPELRSVLOQ
PDSFGKRMFEALLQVLEGIKPNESQVGLPYSELINKELMILPNSIYDSLHNSHQ
IDMDT"
complement(6588. .6935)
/gene="yscg"
/complement(6588. .6935)
/gene="yscg"
/codon_start=1
/transl_table=11
/product="yop proteins translocation protein G homolog"
/protein_id="AAC62548.1"
/db_xref="PID:g2996225"
/db_xref="GI:2996225"
/translation="MKYKLVNLIATETALIGTGNHYHEANCAIEMWLKGEFAVOLI
RSLSLNRGDYASALQOGKRLAPDLEPMALCEYRLGLSALSLRNLRARSQDRI
QTEVNGMRBQKT"
complement(6937. .7200)
/gene="yscf"
/complement(6937. .7200)
/gene="yscf"
/codon_start=1
/transl_table=11
/product="yop proteins translocation protein F homolog"
/protein_id="AAC62549.1"
/db_xref="PID:g2996226"
/db_xref="GI:2996226"
/translation="MSNFSGFTKGTIDIALDANAQTLTKRPADANKAVNDSIALDK
PDNPALLADLQSHINMSVYIININSTIVSMKQMLGILQKFP"
complement(7201. .7401)
/gene="ysce"
/complement(7201. .7401)
/gene="ysce"
/codon_start=1
/transl_table=11
/product="yop proteins translocation protein E homolog"
/protein_id="AAC62550.1"
/db_xref="PID:g2996227"
/db_xref="GI:2996227"
/translation="MTQLEQLHNQETVNSTIQLEMALTKLKKDMNRGDAAQOYVW
QRESKALESAIAlIHVAGDLK"
complement(7398. .8657)

```



8487 DPKRDSVFANRVITTDIELLKILAFLEPEDALIKGCHIDNOLQNGIKR 8536  
296 VKFELESSPTQWELRAFMAMHESFLADRIDDLKVIYDSNMHGDA 345  
|||||  
8537 VKFELESSPTQWELRAFMAMHESFLADRIDDLKVIYDSNMHGDA 8586  
346 SKUREELAEITAEIKIYSVIOAEINKHLSSGTTINHDKSINLMDKNLYG 395  
|||||  
8587 SKUREELAEITAEIKIYSVIOAEINKHLSSGTTINHDKSINLMDKNLYG 8636  
396 YTPDEIRKASAEYKILEKMPOTTIOVDGSEKKIYSINDFLGSEKKRGAL 445  
|||||  
8637 YTPDEIRKASAEYKILEKMPOTTIOVDGSEKKIYSINDFLGSEKKRGAL 8686  
446 GNLKNSYSYKNNNELSHFATCSDKSRPLNDLVSQKTTQLSIDTSRPN 495  
|||||  
8687 GNLKNSYSYKNNNELSHFATCSDKSRPLNDLVSQKTTQLSIDTSRPN 8736  
496 AIEALNRFIOKYDSVMORLIDDTSGK 521  
|||||  
8737 AIEALNRFIOKYDSVMORLIDDTSGK 8762  
seq\_name: gb\_baz:AF074612  
seq\_documentation\_block:  
LOCUS AF074612 70559 bp DNA circular BCT 03-NOV-1998  
DEFINITION Yersinia pestis plasmid pCD1, complete plasmid sequence.  
ACCESSION AF074612  
NID 93822037  
VERSION AF074612.1 GI:3822037  
KEYWORDS  
SOURCE Yersinia pestis.  
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
Yersinia.  
Yersinia pestis.  
1 (bases 1 to 70559)  
Perry,R.D., Straley,S.C., Fetherston,J.D., Rose,D.J., Gregor,J. and  
Blattner,F.R.  
DNA sequencing and analysis of the low-Ca2+-response plasmid pCD1  
of Yersinia pestis KIM5  
Infect. Immun. 66 (10), 4611-4623 (1998)  
JOURNAL 98427122  
MEDLINE 2 (bases 1 to 70559)  
REFERENCE Perry,R.D., Straley,S.C., Fetherston,J.D., Rose,D.J., Gregor,J. and  
AUTHORS Blattner,F.R.  
TITLE Direct Submission  
JOURNAL Submitted (25-JUN-1998) Microbiology and Immunology, University of  
Kentucky, MS415 Medical Center, Lexington, KY 40536-0084, USA  
COMMENT Unique identifiers, starting at Y0001, are assigned to each gene of  
pCD1.  
FEATURES  
source location/Qualifiers  
1. 70559  
/organism="Yersinia pestis"  
/plasmid="pCD1"  
/strain="KIM5"  
/db\_xref="taxon:632"  
57. 368  
/gene="Y0001"  
57. 368  
/gene="Y0001"  
/note="0103: 43 pct identical (0 gaps) to 100 residues of  
an approx. 200 aa protein GENPEPT: g1|537126,orf\_0198  
Escherichia coli"  
/codon\_start=1  
/transl\_table=11  
/product="unknown"  
/protein\_id="AAC69758.1"  
/db\_xref="PID:93822038"  
/db\_xref="GI:3822038"  
/translation="MHQOSRGAAGSRTLSLMROSGYNNVRLARIRARECLASROP  
GKPRYRGEREVSLASPLDKROFKPSEPRKVMSGTISTIKVNGWCYIALVLDLTSFH  
W"  
665. 1033  
gene

/gene="nuc"  
/note="Y0002"  
665. 1033  
/gene="nuc"  
/codon\_start=1  
/transl\_table=11  
/product="endonuclease"  
/protein\_id="AAC69759.1"  
/db\_xref="PID:93822039"  
/db\_xref="GI:3822039"  
/translation="MDTKLQHTPIGTVWDYRPVNTKSGGRRLRCPDFVIHYRDDL  
VNAGIPVRTVNSFKALHDKVIIVDGKNTQWGSFNSQAVOSNSENVLIWGDFTVVO  
AYIQWOSRNMKGIDWRSSY"  
1171. 1425  
/gene="repB"  
/note="Y0003"  
1171. 1425  
/gene="repB"  
/codon\_start=1  
/transl\_table=11  
/product="replication protein"  
/protein\_id="AAC69760.1"  
/db\_xref="PID:93822040"  
/db\_xref="GI:3822040"  
/translation="MSQIENAVTSSSKRAYRKNPLTGAERQMSVSRRKETKHAINV  
FIQNDLKNELLOLEDGSGLTQTEMIERWIQEKKARTVAA"  
complement(1560. 1649)  
/note="antisense RNA"  
/product="copA"  
1667. 1741  
/gene="tap"  
/note="Y0004"  
1667. 1741  
/gene="tap"  
/codon\_start=1  
/transl\_table=11  
/product="repa translation protein"  
/protein\_id="AAC69761.1"  
/db\_xref="PID:93822041"  
/db\_xref="GI:3822041"  
/translation="MFRKVYLLRLLLPCNISAGRCD"  
1734. 2600  
/gene="repa"  
/note="Y0005"  
1734. 2600  
/gene="repa"  
/codon\_start=1  
/transl\_table=11  
/product="replication protein"  
/protein\_id="AAC69762.1"  
/db\_xref="PID:93822042"  
/db\_xref="GI:3822042"  
/translation="WTNHOALFTHHYROVKNPDEFTPESEKKTLPFCRKLMAKEGF  
TSRDFSMHVAFAKSLSLHRMPPLRLRAIDALOGKCFHYDPLANFIORSITLAI  
ECGLATESKSGNLSTYRATRALRELSGLITQTEYDPOIGCNIPDTFTPALFSA  
LDVSPVAVAARBSRVEWENOOREKORLPLEMBELLIKAMRFVRRERSYOTERKAH  
GLKRRARQVDYDTRDIEALVNRQLTREINBGRFVGDLAVRREKARVVERMLSR  
NNNTRLATGAT"  
1741. 2935  
/note="oriR"  
complement(3427. 3645)  
/gene="Y0006"  
complement(3427. 3645)  
/gene="Y0006"  
/note="oriS (f72): 42 pct identical (0 gaps) to 33  
residues of an approx. 216 aa protein GENPEPT: g1|2055297,  
Imp2 Xenopus laevis"  
/codon\_start=1  
/transl\_table=11  
/product="unknown"  
/protein\_id="AAC69763.1"  
/db\_xref="PID:93822043"  
/db\_xref="GI:3822043"

gene /translation="MRSPLAYGCVYTVNVYVYOLHNDHKSAKHKRLHYVLTGQASS  
ADPVFSPHRCMAKPADININSVDNMRR"  
4758. .5186  
/gene="y0008"  
CDS 4758. .5186  
/gene="y0008"  
/note="orf7 (o142): 31 pct identical (1 gap) to 48  
residues of an approx. 104 aa protein GENPEPT: g112149940,  
orf1 p. syringae"  
/codon\_start=1  
/transl\_table=11  
/product="unknown"  
/protein\_id="AAC69764.1"  
/db\_xref="PID:g3822044"  
/db\_xref="GI:3822044"  
/translation="MINTFTLLPKIASHFGDLKLSODEYGLCEILLDRVYMLRA  
DEILNRLTLGLPILGSGPEARSAASQLFFCYISINLNKDGPCFAMSEELGIAPKHL  
SIDELNENVSRIANFYDMLSLVSLPAETAPSYISYSIG"  
5204. .7402  
/gene="ypka"  
/note="y0009"  
CDS 5204. .7402  
/gene="ypka"  
/codon\_start=1  
/transl\_table=11  
/product="secreted protein kinase"  
/protein\_id="AAC69765.1"  
/db\_xref="PID:g3822045"  
/db\_xref="GI:3822045"  
/translation="MKSQKIMGTMPSPISLAKAHERISOHMVPGEINIGKRYRBI  
DNQVLRNHSFSLPREGVKIFSGKMENFSLARLDTOLHAAOCTTQSOELRSDPN  
ALSNLGAKPOTELPLGKGEPLSGAPDLEGMYALTDFABESNISTEIKDKQRL  
VAKIERSIAEGHFALEAKYKLYTAGKHPNLAVNAGAAVAVYGRKKEALLMDVD  
GMCSOTLRLADSMKQKINSEAWGTIKFIHRLIDVYNHAKAGVYNDLTKPNV  
VEPRASGEVYIDLGLHSRSGEQKGFESFKPELGVNGLGASEKSDVLYVSTLH  
CIEGFEKNPEIKNOGLRFTSEPAHVMDENGYPIHFGIAGVETAYTFPIIDILVS  
ADSRPDSNARLHEDISGTIDDESAKOILKDTLGEMSLSTDVRIIPKRIEISD  
LRLHLSAATKOLDMGVSLSDITDMLVALDKERGGVDKDLKFSNLSLTIVYI  
EDYVKGREGDTKNSSTESVYHRSNMLSTVEPSLQRIKHLDQTHSFSIDIGLVAH  
KHLETLLEVLYLSOOGQPVSELYGFLNLTFAKLTLSQNLTLQOQOSESAAQCSI  
LNRSGSMADVAROSLQRFDSIQPVKFGTEOYTAIHRQMAHAALITQEVSEFTDD  
MRNFYDISPILQLGRSSLMDEHLEVRKELBELTIAERLNRLEEMW"  
7798. .8664  
/gene="yopJ"  
/note="y0010"  
CDS 7798. .8664  
/gene="yopJ"  
/codon\_start=1  
/transl\_table=11  
/product="targeted effector protein"  
/protein\_id="AAC69766.1"  
/db\_xref="PID:g3822046"  
/db\_xref="GI:3822046"  
/translation="MIGPISQINISGSLSEKENSLSISNELKNITTOLETDISGSM  
FKHNSRMDEVMPALVIOANKYPMENINLYSPDISLEIKNYIENGVRSRRFIN  
MGBGGIHFSYIDKHNKTSILTFEPANFSGPMLAIRITAIERQLDLPCHFSM  
VEDIDRSSSECGIFSLALAKLYIRDSILKTHEINIGILSDGNPPLPHDLIDYL  
PVYFKHTQCKRLNLTNPQGVGVVKKNETIYVNRDNKNSIVDGKELSVYHK  
KRIAEKYLTKV"  
9858. .10064  
/gene="y0012"  
/note="y0012"  
CDS 9858. .10064  
/gene="y0012"  
/note="o68: 45 pct identical (0 gaps) to 68 residues of an  
approx. 560 aa protein GENPEPT: g112182435, hypothetical"  
/codon\_start=1  
/transl\_table=11  
/product="unknown"  
/protein\_id="AAC69767.1"  
/db\_xref="PID:g3822047"  
/db\_xref="GI:3822047"  
/translation="MAVGRKNMLFAGSLRAGORMASISLETAKLNGHDPYVWLVDY  
LTRLPTWNSQLNALLPYAENRS"

gene 10347. .11753  
/gene="yopH"  
/note="y0013"  
CDS 10347. .11753  
/gene="yopH"  
/codon\_start=1  
alignment\_scores:  
Quality: 1450.00 Length: 326  
Ratio: 4.448 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 99.693.  
alignment\_block:  
US-08-699-716A-2 x AF074612 (2) ..  
Align seg 1/1 to: AF074612 from: 1 to: 70559  
196 MIRAYEQPHQFIEDLEKRVYEOLTGHCSSYLEELQYLNKKNIDISTKY 245  
|||||  
12640 MIRAYEQPHQFIEDLEKRVYEOLTGHCSSYLEELQYLNKKNIDISTKY 12689  
246 DPRKDSEVFANRVITDDIELKILAYFLPEDTILKGGHYNOQLNGIKR 295  
|||||  
12690 DPRKDSEVFANRVITDDIELKILAYFLPEDTILKGGHYNOQLNGIKR 12739  
296 VKPELESSPNTQWELRAFMAYMHFSLADRIDDILKVIYDSNMHGDAR 345  
|||||  
12740 VKPELESSPNTQWELRAFMAYMHFSLADRIDDILKVIYDSNMHGDAR 12789  
346 SKLREELAEFLAEIKIYSVIOAEINKHLSSTGTTINIDKSNIMDKNLYG 395  
|||||  
12790 SKLREELAEFLAEIKIYSVIOAEINKHLSSTGTTINIDKSNIMDKNLYG 12839  
396 YTDEIFKASAEYKILEKPPOTTIOVDSEKKIYSIKDFLGSSEKRRGAL 445  
|||||  
12840 YTDEIFKASAEYKILEKPPOTTIOVDSEKKIYSIKDFLGSSEKRRGAL 12889  
446 GNLKNSIYNNDNNELSHFATCGDSKSRPLNDIVSQKTTQSDITSRFS 495  
|||||  
12890 GNLKNSIYNNDNNELSHFATCGDSKSRPLNDIVSQKTTQSDITSRFS 12939  
496 AIEALNRFIOKYDSVMORLDDTSGK 521  
|||||  
12940 AIEALNRFIOKYDSVMORLDDTSGK 12965  
seq\_name: gb\_ba1:YPRPIVANT  
seq\_documentation block:  
LOCUS YPRPIVANT 981 bp DNA BCT 13-MAR-1997  
DEFINITION Y.pseudotuberculosis v antigen gene.  
ACCESSION X96802  
NID X1405834  
VERSION 96802.1 GI:1405834  
KEYWORDS  
SOURCE Yersinia pseudotuberculosis.  
ORGANISM Yersinia pseudotuberculosis  
Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;  
Yersinia.  
REFERENCE  
1 (bases 1 to 981)  
Roggenkamp, A., Geiger, A.M., Leitritz, L., Kessler, A. and  
Heesemann, U.  
TITLE Passive immunity to infection with Yersinia spp. mediated by  
anti-recombinant v antigen is dependent on polymorphism of v  
antigen  
JOURNAL Infect. Immun. 65 (2), 446-451 (1997)  
MEDLINE 97162308  
REFERENCE 2 (bases 1 to 981)  
Roggenkamp, A.  
AUTHORS Direct Submission  
TITLE Submitted (22-MAR-1996) A. Roggenkamp, Hygiene and Microbiology,  
JOURNAL University of Muenzberg, Josef-Schneider Strasse 2, 97080  
Muenzberg, FRG  
FEATURES  
Location/Qualifiers

source 1. .981 /organism="Yersinia pseudotuberculosis" /sub\_species="type I" /db\_xref="taxon:633" CDS 1. .981 /transl\_table=1 /product="V antigen" /protein\_id="CA65594.1" /db\_xref="PID:e23989" /db\_xref="PID:g1405835" /db\_xref="GI:1405835" /db\_xref="SWISS-PROT:P23994" /translation="MIRAYEQNPQHFIEDLEKRVVOLTGHGSSVLEELVOLVKDKNI DISIKYDPRKSEVFANRVITDDIELKILAFLEPDAILKGGHYDNOQNGIKRVK EFLSESPNTQWELRAFAVAIHFSLTADRIDDLKVIYVDSMNHGDRSKRLRELAEL TAEIKIYVIOAEINKHLSGGTINIHDKSINIMDKNLGYTDEIFKASAEKILIEK MPOTTQEGTEKTIYSIKNFLESEKRTGALGNLKDYSYNKDNNELSHFATTCSDK SRPLNDVSOKTTQSLDITSRFSALNALNRFLOKYDSVMORLLDTSCK"

BASE COUNT 341 a 174 c 191 g 275 t

ORIGIN

alignment\_scores: Quality: 1418.00 Length: 326 Ratio: 4.390 Gaps: 0 Percent Similarity: 99.080 Percent Identity: 96.626

alignment\_block: US-08-699-716a-2 x YPRIVANT (1) ..

Align seg 1/1 to: YPRIVANT from: 1 to: 981

196 MIRAYEQNPQHFIEDLEKRVVOLTGHGSSVLEELVOLVKDKNIDISTKY 245  
|||||  
1 MIRAYEQNPQHFIEDLEKRVVOLTGHGSSVLEELVOLVKDNIDISTKY 50  
246 DPKRSEVFANRVITDDIELKILAFLEPDTILKGGHYDNOQNGIKR 295  
|||||  
51 DPKRSEVFANRVITDDIELKILAFLEPDAILKGGHYDNOQNGIKR 100  
296 VKFLESPNTQWELRAFAVAIHFSLTADRIDDLKVIYVDSMNHGDR 345  
|||||  
101 VKFLESPNTQWELRAFAVAIHFSLTADRIDDLKVIYVDSMNHGDR 150  
346 SKLRELAELTAEIKIYVIOAEINKHLSGGTINIHDKSINIMDKNLGY 395  
|||||  
151 SKLRELAELTAEIKIYVIOAEINKHLSGGTINIHDKSINIMDKNLGY 200  
396 YDEIFKASAEKILIEKMPOTTQEGTEKTIYSIKNFLESEKRTGAL 445  
|||||  
201 YDEIFKASAEKILIEKMPOTTQEGTEKTIYSIKNFLESEKRTGAL 250  
446 GNLKNSYVDSYNKDNNELSHFATTCSDKSRPLNDVSOKTTQSLDITSRFS 495  
|||||  
251 GNLKNSYVDSYNKDNNELSHFATTCSDKSRPLNDVSOKTTQSLDITSRFS 300  
496 ALEALNRFLOKYDSVMORLLDTSCK 521  
|||||  
301 ALEALNRFLOKYDSVMORLLDTSCK 326

seq\_name: gb\_ba1:YEPICRGVHP

seq\_documentation\_block: LOCUS YEPICRGVHP 2201 bp DNA BCT 26-APR-1993 DEFINITION Yersinia pseudotuberculosis V-antigen (lcrG, lcrV, lcrH) genes, complete cds. ACCESSION M57893 NID G155456 VERSION M57893.1 GI:155456 KEYWORDS V-antigen. SOURCE Yersinia pseudotuberculosis (strain YPIII (PIB1) DNA. ORGANISM Yersinia pseudotuberculosis

Bacteria: Proteobacteria: gamma subdivision: Enterobacteriaceae: Yersinia. 1 (bases 1 to 2201) REFERENCE Bergman,T., Hakansson,S., Forsberg,A., Norlander,L., Macellaro,A., Beckman,A., Boelin,I. and Wolf-Watz,H. TITLE Analysis of the V antigen lcrGVH-yopBD Operon of Yersinia pseudotuberculosis: Evidence for a regulatory role of lcrH and lcrV JOURNAL J. Bacteriol. 173, 1607-1616 (1991) MEDLINE 91154114 FEATURES location/Qualifiers source 1. .2201 /organism="Yersinia pseudotuberculosis" /strain="YPIII (PIB1)" /db\_xref="taxon:633" -35\_signal 219. .224 /gene="lcr operon" 219. .258 /gene="lcr operon" 241. .246 /gene="lcr operon" 252. .258 /gene="lcr operon" 264. .551 /gene="lcrG" 264. .551 /gene="lcrG" /db\_xref="PID:g155457" /db\_xref="GI:155457" /translation="MKSSHFDYDKTLKQAEALIASDSRAKILQEMCADIGTPNAV MKTFGRSABEILPAERLDELTKRQERQPOHPYDGKRPKPTMNGQII" 553. .1533 /gene="lcrV" 553. .1533 /gene="lcrV" /db\_xref="lcrv" /codon\_start=1 /transl\_table=1 /evidence=experimental /product="V-antigen" /protein\_id="AA27645.1" /db\_xref="PID:g155458" /db\_xref="GI:155458" /translation="MIRAYEQNPQHFIEDLEKRVVOLTGHGSSVLEELVOLVKDKNI DISIKYDPRKSEVFANRVITDDIELKILAFLEPDAILKGGHYDNOQNGIKRVK EFLSESPNTQWELRAFAVAIHFSLTADRIDDLKVIYVDSMNHGDRSKRLRELAEL TAEIKIYVIOAEINKHLSGGTINIHDKSINIMDKNLGYTDEIFKASAEKILIEK MPOTTQEGTEKTIYSIKNFLESEKRTGALGNLKDYSYNKDNNELSHFATTCSDK SRPLNDVSOKTTQSLDITSRFSALNALNRFLOKYDSVMORLLDTSCK" 1546. .2052 /gene="lcrH" 1546. .2052 /gene="lcrH" /db\_xref="lcrh" /codon\_start=1 /transl\_table=1 /evidence=experimental /product="V-antigen" /protein\_id="AA27646.1" /db\_xref="PID:g155459" /db\_xref="GI:155459" /translation="MOERTTQEQVLAENESFLKGGGTAMLEISSDTLEQYSLAF NOYOSGRYEDAHKVFQALCVLHDYDSRFFLGACQAMGOYLAIHSYSYCAIMDIK EEPFPHAAECULLQKGLAEAGSFLAQLIADKPEFKELSTRVSSMELATLKLEM EHECVNP"

BASE COUNT 714 a 450 c 485 g 552 t

ORIGIN

alignment\_scores: Quality: 1418.00 Length: 326

Ratio: 4.390 Gaps: 0  
Percent Similarity: 99.080 Percent Identity: 96.626

alignment\_block:  
US-08-699-716A-2 x YEP\_LCRG\_VHP (1) ..

Align seg 1/1 to: YEP\_LCRG\_VHP from: 1 to: 2201

```
196 MIRAYEQNPHEFTEDELEKVEVQOLTGHGSSVLEELVOLVYDKNIDISIKY 245
185 MIRAYEQNPHEFTEDELEKVEVQOLTGHGSSVLEELVOLVYDKNIDISIKY 234
246 DPKKDSFVFNRYITDDIELLKIILAFEPEDTILKGGHNDOLQNCIKR 235
235 DPKKDSFVFNRYITDDIELLKIILAFEPEDAILKGGHNDOLQNCIKR 284
296 VKFEFLSSPTQWELRAFMAVHPSLTADRIDDDILKVIYDSNNHHGDAR 345
285 VKFEFLSSPTQWELRAFMAVHPSLTADRIDDDILKVIYDSNNHHGDAR 334
346 SKLREELAEITAEIKIYSVIOAEINKHLSSSGTINIHDKSINLMDKNLYG 395
335 SKLREELAEITAEIKIYSVIOAEINKHLSSSGTINIHDKSINLMDKNLYG 384
396 YTEDEIFKASAEKYLEKMPQTTIYDGSSEKKIYSIKDFIGSEKRTGAL 445
385 YTEDEIFKASAEKYLEKMPQTTIOEGTEKKIYSINFESEKRRGAL 434
446 GNLKNSYSYKKNDELSHFATTCSDKSRPLNDIVSOKTTLSDITSFNS 495
435 GNLKDSYSYKKNDELSHFATTCSDKSRPLNDIVSOKTTLSDITSFNS 484
496 AIEALNRFIOKYDSVMQRLDDTSKG 521
485 AIEALNRFIOKYDSVMQRLDDTSKG 510
```

seq\_name: gb\_ba2:AF102990

seq\_documentation\_block:

LOCUS AF102990 69673 bp DNA circular BCT 18-MAY-1999  
DEFINITION Yersinia enterocolitica plasmid pYve227, complete sequence.  
ACCESSION AF102990 AF054978 AF054979 AF054980 AF054981 AF080156 Z69926 L06216  
AF080155 AF021264 AF050104 AF054977 AF033863 U02499 U08019 U21297  
M22781 U08222 Z18539 U94827  
NTD 94324323  
VERSION AF102990.1 GI:4324323  
KEYWORDS  
SOURCE  
ORGANISM  
Yersinia enterocolitica.  
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
Yersinia.  
REFERENCE  
AUTHORS 1 (bases 32056 to 32871)  
Cornelis, G., Sluiter, C., de Rouvroit, C. L. and Michiels, T.  
TITLE Homology between virF, the transcriptional activator of the  
Yersinia virulence regulon, and Avc, the Escherichia coli  
arabinose operon regulator  
J. Bacteriol. 171 (1), 254-262 (1989)  
JOURNAL MEDLINE 89123026  
AUTHORS 2 (bases 15074 to 15994)  
Hakansson, S., Bergman, T., Vanooteghem, J. C., Cornelis, G. and  
Wolf-Watz, H.  
TITLE YopB and YopD constitute a novel class of Yersinia yop proteins  
JOURNAL MEDLINE 93114907  
AUTHORS 3 (bases 45325 to 45717)  
Wattiau, P. and Cornelis, G. R.  
TITLE SycE, a chaperone-like protein of Yersinia enterocolitica involved  
in Ose secretion of YopE  
Mol. Microbiol. 8 (1), 123-131 (1993)  
JOURNAL MEDLINE 93268087  
AUTHORS 4 (bases 23950 to 25269)  
Woestyn, S., Allaoui, A., Wattiau, P. and Cornelis, G. R.  
TITLE YscN, the putative energizer of the Yersinia yop secretion

\* yscN, the putative energizer of the Yersinia yop secretion

JOURNAL MEDLINE machinery  
AUTHORS J. Bacteriol. 176 (6), 1561-1569 (1994)  
TITLE 94179088  
5 (bases 29897 to 30961)  
Allaoui, A., Woestyn, S., Sluiter, C. and Cornelis, G. R.  
TITLE yscN, a Yersinia enterocolitica inner membrane protein involved in  
yop secretion  
J. Bacteriol. 176 (15), 4534-4542 (1994)  
JOURNAL MEDLINE 94321323  
AUTHORS 6 (bases 43292 to 43717)  
Wattiau, P., Berlier, B., Deslee, P., Michiels, T. and Cornelis, G. R.  
TITLE Individual chaperones required for yop secretion by Yersinia  
proc. Natl. Acad. Sci. U.S.A. 91 (22), 10493-10497 (1994)  
JOURNAL MEDLINE 95024141  
AUTHORS 7 (bases 31537 to 31932)  
Allaoui, A., Scheen, R., Lambert de Rouvroit, C. and Cornelis, G. R.  
TITLE VirG, a Yersinia enterocolitica lipoprotein involved in Ca<sup>2+</sup>  
dependency, is related to exsB of Pseudomonas aeruginosa  
J. Bacteriol. 177 (15), 4230-4237 (1995)  
JOURNAL MEDLINE 95362644  
AUTHORS 8 (bases 53135 to 53497)  
Stainier, I., Iriarte, M. and Cornelis, G. R.  
TITLE yscM1 and yscM2, two Yersinia enterocolitica proteins causing  
downregulation of yop transcription  
Mol. Microbiol. 26 (4), 833-843 (1997)  
JOURNAL MEDLINE 98087284  
AUTHORS 9 (bases 18681 to 18968)  
Boyd, A. P., Sory, M. P., Iriarte, M. and Cornelis, G. R.  
TITLE Heparin interferes with translocation of yop proteins into HeLa  
cells and binds to LcrG, a regulatory component of the Yersinia yop  
apparatus  
Mol. Microbiol. 27 (2), 425-436 (1998)  
JOURNAL MEDLINE 98143428  
AUTHORS 10 (bases 22612 to 22890)  
Iriarte, M., Sory, M. P., Boland, A., Boyd, A. P., Mills, S. D.,  
Lambert, I. and Cornelis, G. R.  
TITLE Tyda, a protein involved in control of Yop release and in  
translocation of Yersinia yop effectors  
EMBO J. 17 (7), 1907-1918 (1998)  
JOURNAL MEDLINE 98190073  
AUTHORS 11 (bases 1 to 69673)  
Iriarte, M., Lambert, I., Kerbourch, C. and Cornelis, G. R.  
TITLE Detailed genetic map of the pYve227 plasmid of Yersinia  
enterocolitica serotype O:9  
Unpublished  
JOURNAL MEDLINE 12 (bases 1 to 69673)  
AUTHORS Iriarte, M., Lambert, I., Kerbourch, C. and Cornelis, G. R.  
TITLE Direct Submission  
Submitted (30-OCT-1998) Microbial Pathogenesis Unit, Christian de  
Duve Institute of Cellular Pathology and Faculte de Medecine,  
Universite Catholique de Louvain, Av. Hippocrate 74.49, Brussels  
1200, Belgium  
COMMENT On May 18, 1999 this sequence version replaced gi:3411136  
gi:3411135 gi:3411139 gi:3411137 gi:3603360 gi:1524380 gi:155536  
gi:3603357 gi:2738986 gi:2944089 gi:3411138 gi:3108216 gi:437201  
gi:497215 gi:885971 gi:155526 gi:483322 gi:48593 gi:2735421.  
FEATURES  
source  
1. 69673  
/organism="Yersinia enterocolitica"  
/strain="W22703"  
/serotype="O:9"  
/db\_xref="taxon:630"  
/plasmid="pYve227"  
53..203  
complement(889..1434)  
/note="ORF181"  
/codon\_start=1  
/transl\_table=11  
/product="unknown"  
/protein\_id="AA016801.1"  
/db\_xref="PID:94324324"  
/db\_xref="GI:4324324"  
/translation="MRVVIYOGGYGIFRKNGLDLSMPAIOITCLGLVALSDQHDYLLCA

rep\_origin  
CDS

CDS  
1616..2083  
/note="ORF155"  
/codon\_start=1  
/transl\_table=11  
/product="unknown"  
/protein\_id="AA016802.1"  
/db\_xref="PID:g4324325"  
/db\_xref="GI:4324325"  
/translation="MINTFTTELLQRIASHFGDLKISODEYGLCELLINDRVYIMTRA  
DEILNRLTLGLPIGSGEPASQLPFCYSINALNKDPCFAMSEELGLIAFKIL  
SLGELNENVSKEIANFYDWLSLVSLEPAFTQELPPTOSTOSTOSTQSVKMG"  
2091..4280  
/gene="yopO"  
2091..4280  
/gene="yopO"  
/codon\_start=1  
/transl\_table=11  
/product="protein kinase YopO"  
/protein\_id="AA016803.1"  
/db\_xref="PID:g4324326"  
/db\_xref="GI:4324326"  
/translation="MKIMGTMPSPISLAKAHERISQHMNPVGLINIGKRRYIINO  
VLRNPHSGFSLFREGVGIIFSGKMFNSIARNLLETILHAOKTISOELRSDIPNAL  
NLFGAPQTEPLPIGKGPGLSGAPDLGEMVAREKRAEESHSIETEDNORIVAK  
IERSIABGHLFAELERAKHITKTRAGKRPANLANYHGAIVYGNKEKALMLDEVDGR  
CSQTLRLADSMKQKGINSEAYWGTIKFIARLLDVLNHLAKAIVANDIKPGVAED  
RAGEPVYIDGLGHSRGEQPKFTSEFKAPLELGVNLSEKSDVLELVSTLHGE  
GPEKDEPIKPNQGLRFTTSEPAHYMDENGPYIHPGJAGVETATRPITDILGASD  
RPSNEARLHEFLSDGTIDESAQKQILKDTLGTGMSPLSTDVRIITPKKRLRELSDLR  
THLSAATKQIDMGVYLSDLDTMLVTLAKAREGCVKDKQLSPNSILKTYSTEDY  
VKREGVTKSSSAVSPPYHRSNEMLSIAEPSLQRIQHLDTQHSFSGISLVRAHKIL  
ETILEVYLVISPOGQPVSEYISFLNRLAKAVTLSQLDTLQQQDSAKAQSLILN  
RSGSMADVARQSLQREDSTRPVYKFTGEQYTAIHRQMAHAATLLOEVESEFTDDMN  
FTADSIPLRLSSRLIDELHVEQREKRELITTAERLNLEREM"  
4485..4760  
/note="ORF91A: unknown function"  
/codon\_start=1  
/transl\_table=11  
/product="unknown"  
/protein\_id="AA016804.1"  
/db\_xref="PID:g4324327"  
/db\_xref="GI:4324327"  
/translation="MISIRKNECPCTLNNGLPVYCGSELAINEGKRVGINNTSALYR  
IMNLGVLRVLPVCSGFTHQRLNVLQGTVPYKIHIAVNSLNKDK"  
4761..5627  
/gene="yopP"  
4761..5627  
/gene="yopP"  
/codon\_start=1  
/transl\_table=11  
/product="Yop effector YopP"  
/protein\_id="AA016805.1"  
/db\_xref="PID:g4324328"  
/db\_xref="GI:4324328"  
/translation="MIGPISQINSFGLSEFETRSLSINBELNIIITOLETDIADGSM  
FHKNYRLDIEVPAVLIQANNKYPEMNLVNSPOLSIIEKNIINGVSSFFIIN  
MGEGLHFSYIDYKINGKTSILFEFVNFNSMPALATISKTALIERYOLPDCHEFM  
VEMDIORSSECGIFSLAKKLTEDRSLIKIHEDNIKIGLSISEMDPLHNKRDPLT  
PVTFYKHTQCKKRLNEVLTNPQCVGVVKKKNTIINRDNKNSIIDGKELSVYRK  
KRIAEYKTLTKY"  
complement(6151..6900)  
/gene="y1pA"  
complement(6151..6900)  
/gene="y1pA"  
/codon\_start=1  
/transl\_table=11  
/product="11poprotein y1pA"  
/protein\_id="AA016806.1"  
/db\_xref="PID:g4324329"  
/db\_xref="GI:4324329"

/translation="MEDDMKNNKLIAITAVLSSVLSCGCGAMSTAIRKRLNEVKTQ  
MSEFTILERSOQTVTLQIKNTSDKNNMLGLAPITAYADKGYTTSPEDAHYWIOA  
NVLKADKMRDREAEGLTSGOYGGAALAGAITGYSNSAGASIGVYAGLVGMV  
ADAMVEDINTMTVDIOISKTPTPIQTDNVALKQGTSSKYQVTSQCNKKQYQTR  
VSSANKVNLKEFEAQPVLEDLAKSIANIL"  
complement(7370..7918)  
/gene="yopO"  
complement(7370..7918)  
/gene="yopO"  
/codon\_start=1  
/transl\_table=11  
/product="yopO"  
/protein\_id="AA016807.1"  
/db\_xref="PID:g4324330"  
/db\_xref="GI:4324330"  
/translation="MTIKDAYNMRALCTALEGAPDTIINTSKENNSYICATAHLR  
TDVCSLVNRVGIPLVSGSLTLEELMVAVGIIYRLYEMQVSDIDTNKRLPNNSD  
FGLVPSYLDGIDIGVITGKKDSKGNIELDPRKNSLLENDIDIKRYLYDENFHRFCIML  
IISKESEELRESQDQKIMG"  
8439..9407  
gene  
alignment\_scores:  
Quality: 1399.00 Length: 323  
Ratio: 4.372 Gaps: 0  
Percent Similarity: 99.071 Percent Identity: 95.356  
alignment\_block:  
US-08-699-716A-2 x AF102990/rev (4) ..  
Align seg 1/1 to reverse of: AF102990 from: 1 to: 69673  
196 MTRAEQNPQHFIEDLEKRVVQLTGHSSVYLEELVOLVKQNDISTIKY 245  
|||||  
6226 MTRAEQNPQHFIEDLEKRVQQLTGHSSVYLEELVOLVKQNDISTIKY 6177  
|||||  
246 DPRKSEYFANRVITDIELLKKIAYFLPEDTILKGGHYDQNLGKIKR 295  
||:|||||  
6176 DPKDSEVPADRVITDIELLKKIAYFLPEALILKGGHYDQNLGKIKR 6127  
|||||  
296 VKPELSSPNTQWELRAFMVNHESLTADRIDDDLKVIYVDSMNHGDKR 345  
|||||  
6126 VKPELSSPNTQWELRAFMVNHESLTADRIDDDLKVIYVDSMNHGDKR 6077  
|||||  
346 SKLRELELTAEIKIYSYQAEINKHLSSSGTIIHKSITLMKNLYG 395  
:|||||  
6076 GLRLELELTAEIKIYSYQAEINKHLSSSGTIIHKSITLMKNLYG 6027  
|||||  
396 YTDEELFKASAEYKILEKMPOTTIOVDSSEKKIYSIKDPLGSENKRTGAL 445  
|||||  
6026 YTDEELFKASAEYKILEKMPOTTIEGETEKIYVIRKFLSEENKRTGAL 5977  
|||||  
446 GNLKNSYSYKDNNELSHFATTCSDKSRLDLVLSQKTTQSLDITSRNS 495  
|||||  
5976 GNLKNSYSYKDNNELSHFATTCSDKSRLDLVLSQKTTQSLDITSRNS 5927  
|||||  
496 ATEALNRFIOKDYDMORLDDT 518  
|||||  
5926 ATEALNRFIOKDYDMORLDDT 5904  
seq\_name: gb\_ba1.ye96PVANT  
seq\_documentation\_block:  
LOCUS YE96PVANT 975 bp DNA BCT 13-MAR-1997  
DEFINITION Y.enteroocollitica v antigen gene, strain Y-96-P.  
ACCESSION X96801  
NID g1405830  
VERSION X96801.1 GI:1405830  
KEYWORDS v antigen.  
SOURCE Yersinia enterocolitica.  
ORGANISM Yersinia enterocolitica  
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
Yersinia.  
REFERENCE 1 (bases 1 to 975)

AUTHORS Rogenkamp,A., Geiger,A.M., Leitritz,L., Kessler,A. and Heesemann,J.  
TITLE Passive immunity to infection with Yersinia spp. mediated by anti-recombinant V antigen is dependent on polymorphism of V antigen  
JOURNAL Infect. Immun. 65 (2), 446-451 (1997)  
MEDLINE 97162308  
REFERENCE 2 (bases 1 to 975)  
AUTHORS Rogenkamp,A.  
TITLE Direct Submission  
JOURNAL Submitted (22-MAR-1996) A. Rogenkamp, Hygiene and Microbiology, University of Wuerzburg, Josef-Schneider Strasse 2, 97080 Wuerzburg, FRG  
FEATURES  
source 1..975  
/organism="Yersinia enterocolitica"  
/db\_xref="taxon:630"  
1..975  
/organism="Yersinia enterocolitica (type 0:3)"  
/strain="y-96-p"  
/db\_xref="taxon:34055"  
1..975  
/codon\_start=1  
/transl\_table=1  
/product="V antigen"  
/protein\_id="CA65593.1"  
/db\_xref="PID:e239988"  
/db\_xref="PID:g1405831"  
/db\_xref="GI:1405831"  
/db\_xref="SPTRMBL:Q56882"  
/translation="MIRAYEQNPQHFIEDLEKRVVOLTGHGSSVLEELVOLVKKNDISIKY  
DISIKYPOKDEVFADRVITDDIELKILAYFLPEDTILKGHHYNOLONGIKRK  
EFLSSPNTQWELRAFMAVHFSLTADRIDDLKIVDSNNHGGARSKLRELAEL  
TAEKITSVIOAEINKLNSDGTINIHDKSINIMDKNLGYTDEIFKASAEYKILK  
MPTTKEGETEKIVSINKNLESENKRTGALGNLKDYSYNKNNELSHGTCSDK  
SRPLNDVSKTTQLSDITSRFNSAIEALNRFIOKYDSVMORLLDTRL"  
BASE COUNT 335 a 177 c 194 g 269 t  
ORIGIN  
alignment\_scores:  
Quality: 1379.00 Length: 323  
Ratio: 4.323 Gaps: 0  
Percent Similarity: 98.762 Percent Identity: 93.808  
alignment\_block:  
US-08-699-716a-2 x YE96PVANT (1) ..  
Align seg 1/1 to: YE96PVANT from: 1 to: 975  
196 MIRAYEQNPQHFIEDLEKRVVOLTGHGSSVLEELVOLVKKNDISIKY 245  
|||||  
1 MIRAYEQNPQHFIEDLEKRVVOLTGHGSSVLEELVOLVQDKKIDISIKY 50  
246 DPKKDESVANRVITDDIELKILAYFLPEDTILKGHHYNOLONGIKR 295  
|||||  
51 DPKKDESVADRVITDDIELKILAYFLPEDAILKGHHYNOLONGIKR 100  
296 VKEFLSSPNTQWELRAFMAVHFSLTADRIDDLKIVDSNNHGGAR 345  
|||||  
101 VKEFLSSPNTQWELRAFMAVHFSLTADRIDDLKIVDSNNHGGAR 150  
346 SKLRELAELTAEKITSVIOAEINKLNSDGTINIHDKSINIMDKNLG 395  
:|||||  
151 SKLRELAELTAEKITSVIOAEINKLNSDGTINIHDKSINIMDKNLG 200  
396 YTDEIFKASAEYKILKMPQTTIQVGSSEKIVSINKDFLGSNNKRTGAL 445  
|||||  
201 YTDEIFKASAEYKILKMPQTTIKGETEKIVSINKNLESENKRTGAL 250  
446 GNKKNYSYNKNNELSHGTCSDKSRPLNDVSKTTQLSDITSRFNS 495  
|||||  
251 GNKKNYSYNKNNELSHGTCSDKSRPLNDVSKTTQLSDITSRFNS 300

496 AIEALNRFIOKYDSVMORLLDPT 518  
|||||  
301 AIEALNRFIOKYDSVMORLLDPT 323  
seq\_name: gb\_bai:YE808VANT  
seq\_documentation\_block:  
LOCUS YE808VANT 975 bp DNA BCT 13-MAR-1997  
DEFINITION Y. enterocolitica V antigen gene, strain 8081.  
ACCESSION X96798  
NID G1405828  
VERSION X96798.1 GI:1405828  
KEYWORDS V antigen.  
SOURCE Yersinia enterocolitica.  
ORGANISM Yersinia enterocolitica  
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
Yersinia.  
REFERENCE 1 (bases 1 to 975)  
AUTHORS Rogenkamp,A., Geiger,A.M., Leitritz,L., Kessler,A. and Heesemann,J.  
TITLE Passive immunity to infection with Yersinia spp. mediated by anti-recombinant V antigen is dependent on polymorphism of V antigen  
JOURNAL Infect. Immun. 65 (2), 446-451 (1997)  
MEDLINE 97162308  
REFERENCE 2 (bases 1 to 975)  
AUTHORS Rogenkamp,A.  
TITLE Direct Submission  
JOURNAL Submitted (22-MAR-1996) A. Rogenkamp, Hygiene and Microbiology, University of Wuerzburg, Josef-Schneider Strasse 2, 97080 Wuerzburg, FRG  
FEATURES  
source 1..975  
/organism="Yersinia enterocolitica"  
/db\_xref="taxon:630"  
1..975  
/organism="Yersinia enterocolitica (type 0:8)"  
/db\_xref="8081"  
/strain="8081"  
/db\_xref="taxon:34054"  
1..975  
/codon\_start=1  
/transl\_table=1  
/product="V antigen"  
/protein\_id="CA65590.1"  
/db\_xref="PID:e239984"  
/db\_xref="PID:g1405829"  
/db\_xref="GI:1405829"  
/db\_xref="SPTRMBL:Q56880"  
/translation="MIRAYEQNPQHFIEDLEKRVVOLTGHGSSVLEELVOLVKKDKKI  
DISIKYPOKDEVFADRVITDDIELKILAYFLPEDAILKGHHYNOLONGIKRK  
EFLSSPNTQWELRAFMAVHFSLTADRIDDLKIVDSNNHGGARSKLRELAEL  
TAEKITSVIOAEINKLNSDGTINIHDKSINIMDKNLGYTDEIFKASAEYKILK  
MSOTTKEGETEKIVSINKNLESENKRTGALGNLKNYSYNKNNELSHGTCSDK  
SRPLNDVSKTTQLSDITSRFNSAIEALNRFIOKYDSVMORLLDTRL"  
BASE COUNT 339 a 170 c 192 g 274 t  
ORIGIN  
alignment\_scores:  
Quality: 1378.00 Length: 323  
Ratio: 4.320 Gaps: 0  
Percent Similarity: 98.762 Percent Identity: 94.427  
alignment\_block:  
US-08-699-716a-2 x YE808VANT (1) ..  
Align seg 1/1 to: YE808VANT from: 1 to: 975  
196 MIRAYEQNPQHFIEDLEKRVVOLTGHGSSVLEELVOLVKKNDISIKY 245  
|||||  
1 MIRAYEQNPQHFIEDLEKRVVOLTGHGSSVLEELVOLVQDKKIDISIKY 50



THIS PAGE BLANK (USPTO)



OM ef: US-08-699-716a-2 to: EST:\* out-format : pfs

Date: Nov 27, 1999 6:49 PM

About: Results were produced by the Gencore software, version 4.5,  
Copyright (c) 1993-1998 Compugen Ltd.

#### Command line parameters:

-MODEL=sw.model -DEV=xlp  
-O=cgnt2.1/USPTO.spool/US08699716/runat\_27111999.151715.6579/app\_query.fasta.1  
-DB=EST -OFMT=fastap -SUFFIX=est -GAPOP=11.000 -GAPEXT=11.000  
-MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000 -GAPOP=4.500  
-OCAPEXT=0.050 -START=1 -DTRANS -MATRIX=pm150  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR\_SCORE=escore  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=stat -USER=US08699716  
-NCPU=6 -ICPU=3 -WAIT -DM\_MEM=28000 -PM\_MEM=15000 -THREADS=1

#### Search information block:

Query: US-08-699-716a-2

Query length: 521

Database: EST:\*

Database sequences: 3032314

Database length: 1171202697

Search time (Sec): 551.310000

#### Score list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
gb_est31:AU061968	3+	83.00	217.91	0.0045	571	AU061968 AU061968 Dictyostellium
gb_est28:AU531330	2+	81.00	211.37	0.0105	621	A1531330 SD02320.5prime SD DRC
gb_est14:AA391919	2+	79.00	208.47	0.0152	442	AA391919 LD10828.5prime LD DRC
gb_est25:AU037853	3+	77.00	200.23	0.0437	587	AU037853 AU037853 Dictyostellium
gb_est19:AA538585	2+	77.00	199.97	0.0452	605	AA538585 LD18150.5prime LD DRC
gb_est11:AA201964	1+	75.00	196.98	0.0664	437	AA201964 LD05053.5prime LD DRC
gb_est24:AU1182515	6+	75.00	196.92	0.0669	440	A1182515 ue21912.x1 Sugaono mod
gb_est14:AA441519	2+	75.00	195.97	0.0756	491	AA441519 LD16485.5prime LD DRC
gb_est14:AA439056	3+	75.00	195.86	0.0766	499	AA439056 LD13501.5prime LD DRC
gb_est18:AA735384	3+	75.00	195.71	0.0781	506	AA735384 LD21220.5prime LD DRC
gb_est21:AA948995	3+	75.00	195.45	0.0807	523	AA948995 LD27722.5prime LD DRC
gb_est20:AA820589	3+	75.00	194.93	0.0864	554	AA820589 LD23394.5prime LD DRC
gb_est11:AA264342	1+	75.00	194.79	0.0879	563	AA264342 LD07582.5prime LD DRC
gb_est19:AA264985	3+	75.00	194.79	0.0879	565	AA264985 LD08662.5prime LD DRC
gb_est25:AU260971	3+	75.00	194.65	0.0895	572	A1260971 LP05026.5prime LP DRC
gb_est25:AU038166	3+	75.00	194.65	0.0895	574	AU038166 AU038166 Dictyostellium
gb_est28:AU511818	3+	75.00	194.43	0.0921	587	A1511818 LD43348.5prime LD DRC
gb_est21:AA940941	2+	75.00	194.38	0.0926	591	AA940941 LD23285.5prime LD DRC
gb_est28:AU511753	1+	75.00	194.25	0.0942	598	A1511753 LD43259.5prime LD DRC
gb_est28:AU511892	2+	75.00	194.12	0.0957	609	A1511892 LD43451.5prime LD DRC
gb_est28:AU514528	3+	75.00	194.04	0.0968	616	A1514528 LD41064.5prime LD DRC
gb_est31:AU061868	2+	74.00	193.90	0.0986	445	AU061868 AU061868 Dictyostellium
gb_est21:AA940652	2+	75.00	193.87	0.0989	627	AA940652 LD22181.5prime LD DRC
gb_est28:AU518356	1+	75.00	193.87	0.0989	625	A1518356 LD37843.5prime LD DRC
gb_est23:AU1133907	2+	75.00	193.83	0.0994	630	A1133907 GH10944.5prime GH DRC
gb_est11:AA264956	1+	75.00	193.31	0.1062	668	AA264956 LD08665.5prime LD DRC
gb_est19:AA263871	1+	75.00	193.27	0.1068	669	AA263871 LD07129.5prime LD DRC
gb_est11:AA264453	1+	75.00	192.79	0.1136	708	AA264453 LD08943.5prime LD DRC
gb_est21:AA979134	1+	75.00	192.79	0.1136	708	AA979134 LD33408.5prime LD DRC
gb_est21:AA941922	3+	75.00	192.75	0.1142	713	AA941922 LD27265.5prime LD DRC
gb_est21:AA942159	3+	75.00	192.61	0.1163	727	AA942159 LD26225.5prime LD DRC
gb_est21:AU039832	1+	71.00	191.51	0.1339	213	AU039832 AU039832 Dictyostellium
gb_est31:AU061961	2+	72.00	190.96	0.1436	319	AU061961 AU061961 Dictyostellium
gb_est27:AU456613	1+	72.00	185.94	0.2733	572	A1456613 LD36691.5prime LD DRC
gb_est33:AU074622	2+	69.00	184.79	0.3171	240	AU074622 AU074622 Dictyostellium
gb_est33:AU045540	3+	70.00	183.08	0.3947	407	AA045540 UI-M-BH-aks-03-0-D
gb_est15:AA048400	4+	70.00	183.02	0.3979	408	AA048400 nfi0e04.s1 NCI CGAP J
gb_est13:AU731994	1+	69.00	179.33	0.5867	416	M73296 EST00444 fetal brain, S
gb_est28:AU503453	3+	69.00	179.33	0.6385	450	A1731994 BNLGH112265 six-day C
gb_est28:AU503453	3+	69.00	177.99	0.7577	544	A1503453 vq86b11.x1 Knowles SC
gb_est23:AU761030	3+	69.00	177.66	0.7914	546	A1761030 w10d07.x1 NCI CGAP C
gb_est23:AU037196	3+	68.00	177.60	0.7970	393	AU037196 Dictyostellium
gb_est33:AU722975	1+	68.00	177.53	0.8038	393	A1722975 fc28b02.y1 Zebrafish
gb_est31:AU060841	6+	66.00	177.00	0.8610	215	AU060841 AU060841 Dictyostellium

gb\_est31:AU060790 3+ 69.00 176.67 0.8981 612 AU060790 AU060790 Dictyostellium

seq\_name: gb\_est31:AU061968

seq\_documentation\_block:

LOCUS AU061968 571 bp mRNA EST 20-MAY-1999  
DEFINITION AU061968 Dictyostellium discoideum SL (H.Urushihara) Dictyostellium  
discoideum cDNA clone SLG812, mRNA sequence.

ACCESSION AU061968

VERSION AU061968.1

KEYWORDS EST, 4883072

SOURCE AU061968.1

ORGANISM Dictyostellium discoideum.

REFERENCE Dictyostellium discoideum.

1 (bases 1 to 571)

Yoshino, R., Morio, T. and Tanaka, Y.

Developmental cDNA in Dictyostellium discoideum

Unpublished (1997)

On Jun 5, 1998 this sequence version replaced gi:3188908.

COMMENT

Contact: Hideko Urushihara

Institute of Biological Sciences

University of Tsukuba

3-3-10 Ten-noda, Tsukuba, Ibaraki 305, Japan

Email: d402huesakura.cc.tsukuba.ac.jp

PROJECT = Dictyostellium discoideum cDNA project in Japan.

FEATURES

source

1. 571

/organism="Dictyostellium discoideum"

/strain="AX4"

/db\_xref="taxon:44689"

/clone="SLG812"

/clone\_1lb="Dictyostellium discoideum SL (H.Urushihara)"

/dev\_stage="slug"

/dev\_stage="slug"

266 a 84 c 56 g 165 t

BASE COUNT

ORIGIN

alignment\_scores:

Quality: 83.00

Ratio: 2.515

Percent Similarity: 78.571

Percent Identity: 35.714

alignment\_block:

US-08-699-716a-2 x AU061968 (3)

Align seg 1/1 to: AU061968 from: 1 to: 571

3 HHHHHHHHSGCHDDDDKHKRSSVATATFCTTATANA 44

||||||| | : : : : : | : : : : : | : : : : : | : : : : : |

91 HHHHHHHHHHHHLLDQOKKRVIIITTTTIIATPIITTT 132

seq\_name: gb\_est28:AU531330

seq\_documentation\_block:

LOCUS AU531330 621 bp mRNA EST 18-MAR-1999

DEFINITION S002320.5prime SD Drosophila melanogaster Schneider L2 cell culture

port2 Drosophila melanogaster cDNA clone S002320 5prime similar to

L22311: shn FBgn0003336 PID:9845332 SPTREMBL:Q24605, mRNA sequence.

ACCESSION AU531330

VERSION 64445465

KEYWORDS EST, 44445465

SOURCE AU531330.1

ORGANISM Drosophila melanogaster

fruit fly.

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Pterygota; Neoptera; Endopterygota; Diptera; Brachyera;

Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 621)

Harvey, D., Hong, L., Evans-Holm, M., Pendleton, J., Su, C.,

Broksstein, P., Lewis, S. and Rubin, G. M.

BDGP/HMI Drosophila EST Project

TITLE



## ORIGIN

alignment\_scores:                   Quality:    77.00                   Length:    25  
                                  Ratio:    3.500                   Gaps:    0  
Percent Similarity: 88.000       Percent Identity: 52.000

## alignment\_block:

US-08-699-716A-2 x AU037853 (3) ..

Align seg 1/1   to: AU037853   from: 1   to: 587

3 HHHHHHHHSSGHIDDDKMKKRI 27

102 HHHHHHHHSGNHINPQCLLNQI 126

seq\_name: gb\_est19:AA538585

## seq\_documentation\_block:

LOCUS   AA538585           605 bp    mRNA           EST           27-NOV-1998

DEFINITION   LD18150.5prime LD Drosophila melanogaster embryo Bluescript

ACCESSION   Drosophila melanogaster CDNA clone LD18150 5prime, mRNA sequence.

NID   AA538585           92794780

VERSION   AA538585.1   GI:2794780

KEYWORDS   EST.

SOURCE   fruit fly.

ORGANISM   Drosophila melanogaster

REFERENCE   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

AUTHORS   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

1 (bases 1 to 605)

Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C.,

Broksstein,P., Lewis,S. and Rubin,G.M.

BDGP/HMI Drosophila EST Project

Unpublished (1997)

On Jan 19, 1998 this sequence version replaced gi:2285101.

COMMENT

TITLE

JOURNAL

Unpublished (1997)

On Jan 19, 1998 this sequence version replaced gi:2285101.

COMMENT

TITLE

JOURNAL

Unpublished (1997)

On Jan 19, 1998 this sequence version replaced gi:2285101.

COMMENT

TITLE

JOURNAL

Align seg 1/1   to: AA538585   from: 1   to: 605

2 GHHHHHHHSSGHIDDDKMKKISS 29

102 GHAXHHHHHTTSANVDGLDTHNLKLP 129

seq\_name: gb\_est11:AA201964

## seq\_documentation\_block:

LOCUS   AA201964           437 bp    mRNA           EST           29-NOV-1998

DEFINITION   LD05053.5prime LD Drosophila melanogaster embryo Bluescript

ACCESSION   Drosophila melanogaster CDNA clone LD05053 5prime, mRNA sequence.

NID   AA201964           91797767

VERSION   AA201964.1   GI:1797767

KEYWORDS   EST.

SOURCE   fruit fly.

ORGANISM   Drosophila melanogaster

REFERENCE   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

AUTHORS   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

1 (bases 1 to 437)

Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C.,

Broksstein,P., Lewis,S. and Rubin,G.M.

BDGP/HMI Drosophila EST Project

Unpublished (1997)

On May 18, 1995 this sequence version replaced gi:811639.

COMMENT

TITLE

JOURNAL

Unpublished (1997)

On May 18, 1995 this sequence version replaced gi:811639.

COMMENT

TITLE

JOURNAL

Unpublished (1997)

On May 18, 1995 this sequence version replaced gi:811639.

COMMENT

TITLE

JOURNAL

Unpublished (1997)

On May 18, 1995 this sequence version replaced gi:811639.

COMMENT

TITLE

JOURNAL

Align seg 1/1   to: AA201964   from: 1   to: 437

2 GHHHHHHHSSGHIDDDKMKKISS 29

109 GHGHHHHHTTSANVDGLDTHNLKLP 136

seq\_name: gb\_est24:AI182515

## seq\_documentation\_block:

LOCUS   AI182515           440 bp    mRNA           EST           08-OCT-1998

DEFINITION   ue21g12.x1 Sugano mouse kidney mklia Mus musculus cDNA clone

ACCESSION   Drosophila melanogaster CDNA clone LD05053 5prime, mRNA sequence.

NID   AA201964           91797767

VERSION   AA201964.1   GI:1797767

KEYWORDS   EST.

SOURCE   fruit fly.

ORGANISM   Drosophila melanogaster

REFERENCE   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

AUTHORS   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

1 (bases 1 to 437)

Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C.,

Broksstein,P., Lewis,S. and Rubin,G.M.

BDGP/HMI Drosophila EST Project

Unpublished (1997)

On May 18, 1995 this sequence version replaced gi:811639.

COMMENT

TITLE

JOURNAL

Unpublished (1997)

On May 18, 1995 this sequence version replaced gi:811639.

COMMENT

TITLE

JOURNAL

Unpublished (1997)

On May 18, 1995 this sequence version replaced gi:811639.

COMMENT

TITLE

JOURNAL

Unpublished (1997)

On May 18, 1995 this sequence version replaced gi:811639.

COMMENT

TITLE

JOURNAL

IMAGE:1481062 3', mRNA sequence.  
 A182515  
 NID 93733153  
 VERSION A182515.1 GI:3733153  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 440)  
 REFERENCE  
 Authors: Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
 Waterston, R.  
 TITLE The WashU-HM1 Mouse EST Project  
 JOURNAL Unpublished (1996)  
 COMMENT On Jan 14, 1998 this sequence version replaced gi:2754392.

CONTACT: Marra M/Mouse EST Project  
 WashU-HM1 Mouse EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@wustl.edu  
 This clone is available royalty-free through LNL ; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:929418  
 Seq primer: custom primer used  
 High quality sequence stop: 420.

FEATURES  
 Source Location/Qualifiers  
 1..440  
 /organism="Mus musculus"  
 /strain="C57BL"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:1481062"  
 /clone\_1ib="Sugano mouse kidney mk1a"  
 /sex="female"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Organ: kidney; Vector: pME18S-FL3; Site 1: DraIII  
 (CACGTGCG); Site 2: DraIII (CACCATGCG); 1st strand cDNA  
 was primed with an oligo(dT) primer  
 [ATGTGGCTTTTCTTTTCTTTTCTTTT]; double-stranded cDNA was  
 ligated to a DraIII adaptor (TGTGGCTACTG6), digested  
 and cloned into distinct DraIII sites of the pME18S-FL3  
 vector (5' site CACGTGCG, 3' site CACCATGCG). XhoI should  
 be used to isolate the cDNA insert. Size selection was  
 performed to exclude fragments <1.5kb. Library  
 constructed by Dr. Sumio Sugano (University of Tokyo  
 Institute of Medical Science). Custom primers for  
 sequencing: 5' end primer CTGTGCTCTTAAAGCTGCG and 3' end  
 primer CGACCTGACGCTGACGACG."

BASE COUNT 91 a 94 c 104 g 151 t  
 ORIGIN

alignment\_scores:  
 Quality: 75.00 Length: 24  
 Ratio: 3.947 Gaps: 0  
 Percent Similarity: 79.167 Percent Identity: 50.000

alignment\_block:  
 US-08-699-716A-2 x A182515/rev (6) ..

Align seg 1/1 to reverse of: A182515 from: 1 to: 440

3 HHHHHHHHSSGHIDDDDKMKK 26  
 . . . . .  
 136 . . . . .  
 seq\_name: gb\_est14:AA441519

seq\_documentation\_block:  
 LOCUS AA441519 491 bp mRNA EST 28-NOV-1998  
 DEFINITION LD16485.5prime LD Drosophila melanogaster embryo Bluescript  
 Drosophila melanogaster cDNA clone LD16485 5prime, mRNA sequence.  
 ACCESSION AA441519  
 NID 92153397  
 VERSION AA441519.1 GI:2153397  
 KEYWORDS EST.  
 SOURCE fruit fly.  
 ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 Pterygota; Neoptera; Endopterygota; Diptera; Brachyera;  
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 491)  
 REFERENCE  
 Authors: Harvey, D., Hong, L., Evans-Holm, M., Pendleton, J., Su, C.,  
 Brokstein, P., Lewis, S., and Rubin, G.M.  
 TITLE BDGP/HM1 Drosophila EST Project  
 JOURNAL Unpublished (1997)  
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1395398.

CONTACT: Harvey, D.  
 G. M. Rubin-Molecular and Cell Biology  
 University of California Berkeley  
 539 LSA, Berkeley, CA 94720-3200, USA  
 Fax: 510 643 9947  
 Email: http://www.fruitfly.org/EST\_est@fruitfly.berkeley.edu  
 Plate: 164 row: H column: 1  
 High quality sequence stop: 363.

FEATURES  
 Source Location/Qualifiers  
 1..491  
 /organism="Drosophila melanogaster"  
 /db\_xref="BDGP\_EST:BDc1n015696"  
 /db\_xref="taxon:7227"  
 /clone="LD16485"  
 /clone\_1ib="LD Drosophila melanogaster embryo Bluescript"  
 /sex="male and female"  
 /dev\_stage="0 to 24 hours mixed stage embryonic"  
 /lab\_host="SOLR"  
 /note="Organ: embryo; Vector: Bluescript SK; Site 1:  
 EcoRI; Site 2: XhoI; Constructed using Stratagene ZAP-cDNA  
 Synthesis kit. Oligo dT-primed and directionally cloned at  
 EcoRI and XhoI in Bluescript SK(+/-)"

BASE COUNT 104 a 153 c 138 g 94 t  
 ORIGIN

alignment\_scores:  
 Quality: 75.00 Length: 28  
 Ratio: 3.125 Gaps: 0  
 Percent Similarity: 85.714 Percent Identity: 53.571

alignment\_block:  
 US-08-699-716A-2 x AA441519 (2) ..

Align seg 1/1 to: AA441519 from: 1 to: 491

2 GHHHHHHHSSGHIDDDDKMKKISS 29  
 . . . . .  
 100 GHHGHHHHHHTSANVDDGLDTHNLKLP 127

seq\_name: gb\_est14:AA439056

seq\_documentation\_block:  
 LOCUS AA439056 499 bp mRNA EST 28-NOV-1998  
 DEFINITION LD13501.5prime LD Drosophila melanogaster embryo Bluescript  
 Drosophila melanogaster cDNA clone LD13501 5prime, mRNA sequence.  
 ACCESSION AA439056  
 NID 92150937  
 VERSION AA439056.1 GI:2150937  
 KEYWORDS EST.  
 SOURCE fruit fly.  
 ORGANISM Drosophila melanogaster

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Harvey, D., Hong, L., Evans-Holm, M., Pendleton, J., Su, C.,  
Brokstein, P., Lewis, S. and Rubin, G.M.  
BDGP/HMM Drosophila Est Project  
Unpublished (1997)  
On Sep 12, 1996 this sequence version replaced gi:13939447.

Contact: Harvey, D.  
G. M. Rubin-Molecular and Cell Biology  
University of California Berkeley  
539 LSA, Berkeley, CA 94720-3200, USA  
Fax: 510 643 9947  
Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)  
Plate: 135 row: A column: 1  
High quality sequence stop: 427.

FEATURES	Location/Qualifiers
source	1. .499

BASE COUNT	107 a	154 c	145 g	93 t
------------	-------	-------	-------	------

[illegible]

```
alignment_block:
US-08-699-716A-2 x AA439056 (3) ..
Align seg 1/1 to: AA439056 from: 1 to: 499
```

```

2  GHHHHHHHHHSSGHIDDDDKHHMKISS 29
   || |||||:|:::|:| |::
100 GHHGHHHHHHHTSANVDGLDTHNLKLP 127

```

seq_name:	gb_test18:AA735384
seq_documentation_block:	
LOCUS	AA735384 506 bp mRNA
DEFINITION	LD121220 5prime LD Drosophila melanogaster embryo pot2 Drosophila melanogaster CDNA clone LD21220 5prime, mRNA sequence.

NID	92761314
VERSION	AA735384.1
KEYWORDS	EST.
SOURCE	Fruit fly.
ORGANISM	Drosophila melanogaster

REFERENCE 1 (bases 1 to 506)  
AUTHORS Harvey, D., Hong, L., Evans-Holm, M., Pendleton, J., Su, C.,  
Broksstein, P., Lewis, S. and Rubin, G. M.  
TITLE BDGP/HMM Drosophila EST Project  
JOURNAL Unpublished (1997)  
COMMENT On Jan 5, 1998 this sequence version replaced gi:2581344  
Contact: Harvey, D.

G. M. Rubin-Molecular and Cell Biology  
University of California Berkeley  
539 LSA, Berkeley, CA 94720-3200, USA  
Fax: 510 643 9947  
Email: <http://www.frnfly.org/EST>, [est@frnfly.berkeley.edu](mailto:est@frnfly.berkeley.edu)  
Plate: 212 Row: B Column: 393.  
High quality sequence stop: 393.

FEATURES  
SOURCE

Page Count 100 160 110 70 +

```
alignment_scores:
  Quality: 75.00
  Ratio: 3.125
  Gaps: 0
Percent Similarity: 85.714
Percent Identity: 53.571
```

```
alignment_block:
US-08-699-716A-2 x AA735384 (3) ..
```

```

2  GHHHHHHHHHSSGHIDDDDKHKMKISS 29
   ||| |||||:|:::|:|:|:|:|:|:|
86 GHHGHHHHHHHTSANVDGLDTHNLKLP 113

```

seq_name	seq_documentation_block	LOCUS	523 bp	mRNA	EST	25-NOV-1998
seq_name: gb_est21:AA948995		LOCUS	523 bp	mRNA	EST	25-NOV-1998
		DEFINITION	LD27722.5prime	LD Drosophila melanogaster embryo	pot2 Drosophila	
			melanogaster cDNA clone	LD27722 5prime, mRNA sequence.		

VERSION	AA948995.1	GI:3110904
KEYWORDS	EST.	
SOURCE	fruit fly.	
ORGANISM	Drosophila melanogaster	

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 523)	Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C., Brokstein,P., Lewis,S. and Rubin,G.M.	BDGP/HMIT Drosophila EST Project	Unpublished (1997)	On Jan 19, 1998 this sequence version replaced gi:2284655

```

FEATURES
  source      location/Qualifiers
              1..523
              /organism="Drosophila melanogaster"
              /db_xref="taxon:7227"
              /clone="LD27722"
              /clone_lib="LD Drosophila melanogaster embryo POT2"

```

```

seq_name: gb_est20:AA820589
seq_documentation_block:
LOCUS      AA820589          554 bp            mRNA                    EST
DEFINITION LD24394.sprime LD Drosophila melanogaster embryo POT2 Drosophila
melanogaster cDNA clone LD24394 5prime, mRNA sequence.
ACCESSION   AA820589
VERSION     g2890457
KEYWORDS    AA820589.1 GI:2890457
SOURCE      EST.
ORGANISM    fruit fly.
             Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Ephydriidae; Drosophila.
REFERENCE   1 (bases 1 to 554)
AUTHORS     Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C.,
             Brokstein,P., Lewis,S. and Rubin,G.M.
TITLE       BDGP/HHMT Drosophila EST Project
JOURNAL     Unpublished (1997)
COMMENT     On Jan 19, 1998 this sequence version replaced gi:2152541.

BASE COUNT      114 a      165 c      148 g      96 t
ORIGIN

alignment_scores:
Quality:        75.00           Length:         28
Ratio:          3.125           Gaps:          0
Percent Similarity: 85.714       Percent Identity: 53.571

alignment_block:
US-08-699-716A-2 x AA948995 (3) ..

Align seg 1/1 to: AA948995 from: 1 to: 523

      2 GHHNNHHNHHSSGHTDDDKMKRTSS 29
      ||| |||||:::|::| |::|
      100 GHNGHHNHHNHTSANVDGLDTNLKLPRA 127

seq_name: gb_est20:AA820589
seq_documentation_block:
LOCUS      AA820589          554 bp            mRNA                    EST
DEFINITION LD24394.sprime LD Drosophila melanogaster embryo POT2 Drosophila
melanogaster cDNA clone LD24394 5prime, mRNA sequence.
ACCESSION   AA820589
VERSION     g2890457
KEYWORDS    AA820589.1 GI:2890457
SOURCE      EST.
ORGANISM    fruit fly.
             Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Ephydriidae; Drosophila.
REFERENCE   1 (bases 1 to 554)
AUTHORS     Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C.,
             Brokstein,P., Lewis,S. and Rubin,G.M.
TITLE       BDGP/HHMT Drosophila EST Project
JOURNAL     Unpublished (1997)
COMMENT     On Jan 19, 1998 this sequence version replaced gi:2152541.

BASE COUNT      129 a      171 c      156 g      98 t
ORIGIN

FEATURES
Source
  /organism="Drosophila melanogaster"
  /db_xref="taxon:7227"
  /map="11 q13.4-q23.2"
  /clone="LD24394"
  /clone_lib="LD Drosophila melanogaster embryo POT2"
  /sex="male and female"
  /dev_stage="0 to 24 hours mixed stage embryonic"
  /lab_host="Xl1 Blue"
  /note="Organ: embryo; Vector: pOT2; Site_1: EcoRI; Site_2
XhoI; Sized fractionated cDNAs were directly ligated into
pOT2."

```

```

Ratio: 3.125      Gaps: 0
Percent Similarity: 85.714      Percent Identity: 53.571

alignment_block:
US-08-699-716A-2 x AA820589 (3) ..

Align seg 1/1 to: AA820589 from: 1 to: 554

2 GHHNNHHHHSSGHIIDDDKHKRISS 29
||| |||||:::|::|::|::|
101 GHHGHHHHHNTSANNVDCIDTINLKLPA 128

seq_name: gb_est11:AA264342

seq_documentation_block:
LOCUS AA264342 563 bp mRNA EST 29-NOV-1998
DEFINITION LD07582.5prime LD Drosophila melanogaster embryo Bluescript
Drosophila melanogaster cDNA clone LD07582 5prime, mRNA sequence.
ACCESSION AA264342
NID g1900420
VERSION AA264342.1 GI:1900420
KEYWORDS EST.
SOURCE EST.
ORGANISM Drosophila melanogaster
fruit fly.
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachyera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 563)
Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C.,
Brokstein,P., Lewis,S. and Rubin,G.M.
BDGP/HM1 Drosophila EST Project
Unpublished (1997)
On May 8, 1995 this sequence version replaced gi:801550.

TITLE
JOURNAL

FEATURES
source
1..563
/organism="Drosophila melanogaster"
/db_xref="BDGP EST:BDc1n006957"
/db_xref="taxon:7227"
/clone="LD07582"
/clone_1b="LD Drosophila melanogaster embryo Bluescript"
/sex="male and female"
/dev_stage="0 to 24 hours mixed stage embryonic"
/lab_host="SOLR"
/note="Organ: embryo; Vector: Bluescript SK; Site: 1;
ECORI. Site2: XhoI; Constructed using Stratagene Zap-cDNA
Synthesis kit. Oligo dt-primed and directionally cloned at
ECORI and XhoI in Bluescript SK(+/-)"
130 a 175 c 155 g 103 t

BASE COUNT
ORIGIN

alignment_scores:
Quality: 75.00 Length: 28
Ratio: 3.125 Gaps: 0
Percent Similarity: 85.714 Percent Identity: 53.571

alignment_block:
US-08-699-716A-2 x AA264342 (1) ..

Align seg 1/1 to: AA264342 from: 1 to: 563

2 GHHNNHHHHSSGHIIDDDKHKRISS 29
||| |||||:::|::|::|::|
112 GHHGHHHHHNTSANNVDCIDTINLKLPA 139

```

```

2  GHHHHHHHHSSGHIDDDDKHMKISS 29
   ||| |||||:~::~: |~: |~::
100 GHHGHHHHHHHTSANVDGLDTHNKLPA 12

```

11





```

174 GCKLAGKATTDVATVTVSNQ...EMIRAYEONPOHFIEDLEKVEOULT 220
|||||
155 GKLAGKATTDVATVTVSNQSIGIRIRAYEONPOHFIEDLEKVEOULT 204
|||||
221 GGGSSYLEELVOLVKDKNIDISIKYDPRKDSVPANRVITDDIELKTKIL 270
|||||
205 GGGSSYLEELVOLVKDKNIDISIKYDPRKDSVPANRVITDDIELKTKIL 254
|||||
271 AAYLEEDTILKGGHYDNOLONGIKRKYKEFLSSPNTOWELRAFAMVMHFS 320
|||||
255 AAYLEPEDATILKGGHYDNOLONGIKRKYKEFLSSPNTOWELRAFAMVMHFS 304
|||||
321 LTADRIDDDILKVIYDSNMHHODASKRLRELAELTAEIKTISVQAEIN 370
|||||
305 LTADRIDDDILKVIYDSNMHHODASKRLRELAELTAEIKTISVQAEIN 354
|||||
371 KHLSSSGTINIHDKSINLMDKNLYGYTDEIFKASAEKYLEKMPOTTIQ 420
|||||
355 KHLSSSGTINIHDEGINLMDKNLYGYTDEIFKASAEKYLEKMPOTTIQ 404
|||||
421 VDGESEKKIYISINDPLGSENNKRTGALGNLKNYSYNKDNNEISHFATTCSD 470
|||||
405 VDGESEKKIYISINDPLGSENNKRTGALGNLKNYSYNKDNNEISHFATTCSD 454
|||||
471 KRRPLNDIVSOKTTQLOSDITSRFSNATIEALNFIQKYDSVMORLLDDTSG 520
|||||
455 KSRPLNDIVSOKTTQLOSDITSRFSNATIEALNFIQKYDSVMORLLDDTSG 504
|||||
521 K 521
|
505 K 505

seq_name: N_Geneseq_36.T38256

seq_documentation_block:
ID T38256 standard; DNA; 1462 BP.
AC T38256;
DE 28-DEC-1996 (first entry)
Y. pestis F1/V antigen gene fusion.
KW Plaque; vaccine; genetic immunisation; V antigen; lcrv;
KM F1 antigen; calf; ds.
OS Chimeric Yersinia pestis strain GB;
FH Chimeric synthetic.
FT Key
FT Cds
FT location/Qualifiers
FT 8..1450
FT /tag= a
FT /product= F1/V fusion protein
FT misc_feature
FT 452..472
FT /tag= b
FT /note= "bases 452-472 is a sequence derived
FT from PCR primers"
FT
FT
FN WO9628551-A1.
PN 19-SEP-1996.
PF 13-MAR-1996; G00571.
PR 13-MAR-1995; GB-005059.
PR 15-SEP-1995; GB-018946.
PR 05-DEC-1995; GB-024825.
PR (MINA) UK SEC FOR DEBENCE.
PA Bennett AM, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
PI WPI: 96-433824/43.
DR P-PSDB: W01045.
PT Yersinia pestis V antigen and F1 antigen or their protective
PT epitopic parts - useful in vaccine for protection against plague
PS Disclosure; Page 51:55; 96pp; English.
CC A gene fusion (T38256) comprises coding sequences for the Yersinia
CC pestis F1 antigen (see also T38244) (without the signal sequence)
CC and for the Y. pestis V antigen (see also T38243), joined by a
CC sequence encoding a 6-amino acid peptide linker. It was obt. by
CC PCR amplification of the individual genes using primers (see also
CC T38245) based on the genes and including linker sequences. The
CC gene fusion (see also T38249) can be used to produce F1/V fusion
CC protein (W01045) in transformed cells, esp. gut-colonising

```

CC	organisms, to induce an immune response against Y. pestis, the
CC	causative organism of plague.
SQ	Sequence 1462 BP; 476 A; 285 C; 301 G; 400 T;
alignment_scores:	
Quality:	2029.00
Ratio:	4.254
Percent Similarity:	98.963
Percent Identity:	98.133
alignment_block:	
US-08-699-716A-2 x T38256 (2)	..
Align seg 1/1	to: T38256 from: 1 to: 1462
43	NNADLTASTATATLLEPARITLTLYKEGAPITIMDGNIDITELLVGLTL 92
1	SSADLTASTATATLLEPARITLTLYKEGAPITIMDGNIDITELLVGLTL 50
93	GGKGTGTSISVNFDTAAGDPMYLTFTSODGNHNGFTTKYIGKDSRFDI 142
51	GGKGTGTSISVNFDTAAGDPMYLTFTSODGNHNGFTTKYIGKDSRFDI 100
143	SPKVGCEINTVGGDVLVATGSGDFVVSIGSGKGLAAGITDAVTVTVSN 152
101	SPKVGCEINTVGGDVLVATGSGDFVVSIGSGKGLAAGITDAVTVTVSN 150
193	O..EEMIRAYBONPHFTIEDLEKVAEOLITGHGSSVLELVOLVADKNI 239
151	QGSIEGRIRAYEONPHFTIEDLEKVAEOLITGHGSSVLELVOLVADKNI 200
240	DISIKYDPKRDSEVFANRYITDDIELLKKILAYFLPEDITLKGGHYDNL 289
201	DISIKYDPKRDSEVFANRYITDDIELLKKILAYFLPEDITLKGGHYDNL 250
290	QNGIKRVKFELESSEPTQWELRAFMAVMHESLTADRIDDIKVIYDSKN 339
251	QNGIKRVKFELESSEPTQWELRAFMAVMHESLTADRIDDIKVIYDSKN 300
340	HGGDARSKLREELAEITAEIKIYSVIOAEINRKLSSGTINIHDKSINLM 389
301	HGGDARSKLREELAEITAEIKIYSVIOAEINRKLSSGTINIHDKSINLM 350
390	DKNLVYGTDEEIEFKAAEYKILEKMPOTTIOVDSSEKKIYSINDPLGSEN 439
351	DKNLVYGTDEEIEFKAAEYKILEKMPOTTIOVDSSEKKIYSINDPLGSEN 400
440	KRTGALGINKNSYSYNNKDNNELSHEATTSDDSKSRPLNDLVSOQTTOISDI 489
401	KRTGALGINKNSYSYNNKDNNELSHEATTSDDSKSRPLNDLVSOQTTOISDI 450
490	TSRFNSAIEALNRFIOKXDSYVMORLLDDTSRG 521
451	TSRFNSAIEALNRFIOKXDSYVMORLLDDTSRG 482
seq_name:	N_Geneseq_36:T04222
seq_documentation_block:	
ID	T04222 standard; DNA: 1014 BP.
AC	T04222.
DT	18-APR-1996 (first entry)
DE	Partial lcrV (V antigen) gene of Y. pestis.
KW	lcrV: V antigen; virulence; plague; vaccine; epitope; ss.
OS	Yersinia pestis.
FT	key
FT	Location/Qualifiers
FT	cds
FT	1..990
FT	/tag= a
FT	/note= "V antigen"
PN	WO9524475-A1.
PD	14-SEP-1995.
FF	06-MAR-1995; G00481.
PR	08-MAR-1994; GB-004577.
PA	(MINA ) UK SEC FOR DEFENCE.

PI leary SEC, Titball RW, Williamson ED, leary SE;  
DR WPI: 95-328268/42.  
DR P-PSDB: R/9961.  
PT Recombinant DNA expressing Yersinia pestis V antigen - useful in  
PT oral or parenteral vaccines for protection against plague  
PS Claim 6: Page 11-13: 25pp: English.  
CC T04222-23 are DNA sequences (lcry) encoding all or a protective epitopic  
CC part of the mature V protein of Yersinia pestis. The protein was  
CC expressed as a fusion protein with maltose binding protein or  
CC glutathione-S-transferase in 3 different plasmid vectors. Y. pestis is  
CC the highly virulent causative organism of plague in a wide range of  
CC animals, including man. The V antigen (lcry) is an unstable 37.3 kDa  
CC monomeric peptide encoded on the ca. 70 kb lcr plasmid. The V antigen  
CC is postulated to act as a virulence antigen, and transformed  
CC microorganisms contg. recombinant DNA encoding a V antigen protein/  
CC peptide are useful in vaccines to protect against plague.  
SQ Sequence 1014 BP; 346 A; 181 C; 201 G; 286 T;

alignment\_scores:  
Quality: 1443.00 Length: 325  
Ratio: 4.440 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 99.692

alignment\_block:  
US-08-699-716a-2 x T04222 (1) ..  
Align seg 1/1 to: T04222 from: 1 to: 1014

197 IRAYEONPOHFIEDLEKVRVEQDLTGHSVLEELVOLVKKNIDISIKYD 246  
|||||  
5 IRAYEONPOHFIEDLEKVRVEQDLTGHSVLEELVOLVKKNIDISIKYD 54  
247 PRDSEVFARVITDDIELKILAYFLPEDTLKGGHYDNOQNGIKRV 296  
|||||  
55 PRDSEVFARVITDDIELKILAYFLPEDTLKGGHYDNOQNGIKRV 104  
297 KEFLESSPNTQWELRAFMAVMHFSLTADRIDDLIKYIVDSMNHGDARS 346  
|||||  
105 KEFLESSPNTQWELRAFMAVMHFSLTADRIDDLIKYIVDSMNHGDARS 154  
347 KLREBELAELTAEIKIYSVIOAEINKHLSSSGTINIHDKSINLMDKNLYGY 396  
|||||  
155 KLREBELAELTAEIKIYSVIOAEINKHLSSSGTINIHDKSINLMDKNLYGY 204  
397 TDEIFKASAEYKILEKMPOTTIOVDSSEKKIYSIKDPLGSEKRRGALG 446  
|||||  
205 TDEIFKASAEYKILEKMPOTTIOVDSSEKKIYSIKDPLGSEKRRGALG 254  
447 NLKNSYSYKNDNNELSHFATTCSDKSRPLNDIVSOKTTOJSDITSRNSA 496  
|||||  
255 NLKNSYSYKNDNNELSHFATTCSDKSRPLNDIVSOKTTOJSDITSRNSA 304  
497 IEALNRFIOKYDSVMORLLDDTSGK 521  
|||||  
305 IEALNRFIOKYDSVMORLLDDTSGK 329

seq\_name: N\_Geneseq\_36:T38242

seq\_documentation\_block:  
ID T38242 standard; DNA: 1014 BP.  
AC T38242;  
DT 28-DEC-1996 (first entry)  
DE Y. pestis lcrv (V antigen) gene.  
KW Plague; vaccine; genetic immunisation; V antigen; lcrv;  
KW Yersinia pestis strain GB.  
OS Yersinia pestis  
FS key Location/Qualifiers  
FT cds 1..990  
FT misc-feature 1..6  
FT FT /\*tag= a  
FT FT /\*tag= b  
FT FT /note= "vector pMAL-p2 or pMAL-c2-derived bases"

FT mutation 1006  
FT FT /\*tag= C  
FT FT /note= "base 1006 is altered to a T to create a  
FT second in-frame stop codon"  
PN WO9628551-A1.  
PP 19-SEP-1996.  
PP 13-MAR-1996; G00571.  
PP 13-MAR-1995; GB-005059.  
PP 15-SEP-1995; GB-018946.  
PP 05-DEC-1995; GB-024825.  
PA (MINA) UK SEC FOR DEFENCE.  
PI Bennett AM, leary SEC, Oyston PCF, Titball RW, Williamson ED;  
DR WPI: 96-433824/43.  
DR P-PSDB: W01040.  
PT Yersinia pestis V antigen and F1 antigen or their protective  
PT epitopic parts - useful in vaccine for protection against plague  
PS Claim 5: Page 25-28: 98pp: English.  
CC A lcrv gene sequence (T38242) codes for the Yersinia pestis V  
CC antigen (W01040), which is capable of evoking protective immune  
CC responses in animals. The gene was amplified from Y. pestis  
CC DNA by PCR using primers (T38250-51) homologous to the 5' and 3'  
CC ends of the gene. The gene was inserted into vector pMAL-p2,  
CC pMAL-c2 or pGEX-5x-2 (see also T38243) to allow prodn. of  
CC recombinant V antigen for use in vaccines against plague.  
CC Expression in gut-colonising organisms and attenuated Salmonella  
CC typhi allows prodn. of live vaccines. F1/V antigen fusions were  
CC also created (see also T38249 and T38256). The gene can itself be  
CC used in genetic vaccines.  
SQ Sequence 1014 BP; 347 A; 180 C; 201 G; 286 T;

alignment\_scores:  
Quality: 1443.00 Length: 325  
Ratio: 4.440 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 99.692

alignment\_block:  
US-08-699-716a-2 x T38242 (1) ..  
Align seg 1/1 to: T38242 from: 1 to: 1014

197 IRAYEONPOHFIEDLEKVRVEQDLTGHSVLEELVOLVKKNIDISIKYD 246  
|||||  
5 IRAYEONPOHFIEDLEKVRVEQDLTGHSVLEELVOLVKKNIDISIKYD 54  
247 PRDSEVFARVITDDIELKILAYFLPEDTLKGGHYDNOQNGIKRV 296  
|||||  
55 PRDSEVFARVITDDIELKILAYFLPEDTLKGGHYDNOQNGIKRV 104  
297 KEFLESSPNTQWELRAFMAVMHFSLTADRIDDLIKYIVDSMNHGDARS 346  
|||||  
105 KEFLESSPNTQWELRAFMAVMHFSLTADRIDDLIKYIVDSMNHGDARS 154  
347 KLREBELAELTAEIKIYSVIOAEINKHLSSSGTINIHDKSINLMDKNLYGY 396  
|||||  
155 KLREBELAELTAEIKIYSVIOAEINKHLSSSGTINIHDKSINLMDKNLYGY 204  
397 TDEIFKASAEYKILEKMPOTTIOVDSSEKKIYSIKDPLGSEKRRGALG 446  
|||||  
205 TDEIFKASAEYKILEKMPOTTIOVDSSEKKIYSIKDPLGSEKRRGALG 254  
447 NLKNSYSYKNDNNELSHFATTCSDKSRPLNDIVSOKTTOJSDITSRNSA 496  
|||||  
255 NLKNSYSYKNDNNELSHFATTCSDKSRPLNDIVSOKTTOJSDITSRNSA 304  
497 IEALNRFIOKYDSVMORLLDDTSGK 521  
|||||  
305 IEALNRFIOKYDSVMORLLDDTSGK 329

seq\_name: N\_Geneseq\_36:T38243

seq\_documentation\_block:  
ID T38243 standard; DNA: 1014 BP.

```
AC T38243:
DE 28-DEC-1996 (first entry)
DE Y. pestis lcrV (V antigen) gene.
KW Plaque; vaccine; genetic immunisation; V antigen; lcrV;
KW Fl antigen; ds.
OS Yersinia pestis strain GB.
FH Key location/Qualifiers
FT cds 1..990
FT misc-feature /*tag= a
FT /*tag= 1..10
FT /*tag= b
FT /*note= "vector pGEX-5X-2-derived bases"
FT mutation 16
FT /*tag= c
FT /*note= "base 16 is altered to a C from an
FT A to createan EcoRI site"
FT mutation 1006
FT /*tag= d
FT /*note= "base 1006 is altered to a T to create a
FT second in-frame stop codon"
PN W09628551-A1.
PD 19-SEP-1996.
PF 13-MAR-1995: G00571.
PR 13-MAR-1995: GB-005059.
PR 15-SEP-1995: GB-018946.
PR 05-DEC-1995: GB-024825.
PA (MINA ) UK SEC FOR DEFENCE.
PI Bennett AM, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
DR WPI; 96-433824/43.
DR P-PSDB: W01041.
PT Yersinia pestis V antigen and Fl antigen or their protective
PT epitopic parts - useful in vaccine for protection against plague
CC Claim 5: Page 32-35; 98pp; English.
CC A lcrV gene sequence (T38243) codes for the Yersinia pestis V
CC antigen (W01041), which is capable of evoking protective immune
CC responses in animals. The gene was amplified from Y. pestis
CC DNA by PCR using primers (T38251 and T38259) homologous to the 5'
CC and 3' ends of the gene. The gene was inserted into vector
CC pEX-5X-2, PMAL-p2 or PMAL-c2 (see also T38242) to allow prodn. of
CC recombinant V antigen for use in vaccines against plague.
CC Expression in gut-colonising organisms and attenuated Salmonella
CC typhi allows live vaccine prodn. Fl/V antigen fusions were also
CC created (see also T38249 and T38256). The gene can itself be used
CC in genetic vaccines.
SQ Sequence 1014 BP; 343 A; 184 C; 205 G; 282 T;

alignment_scores:
Quality: 1443.00 Length: 325
Ratio: 4.440 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.692

alignment_block:
US-08-699-716A-2 x T38243 (1) ..
Align seg 1/1 to: T38243 from: 1 to: 1014

197 IRAYEONPOHFIEDELEKRVVEOLTGHGSSVLEELVOLVKDKNIDISTKYD 246
|||||
5 IRAYEONPOHFIEDELEKRVVEOLTGHGSSVLEELVOLVKDKNIDISTKYD 54
247 PRKQSEVFNANRVITDDIELLKKILAYFLPEPTILKGGHYDNLONGIKRV 296
|||||
55 PRKQSEVFNANRVITDDIELLKKILAYFLPEPTILKGGHYDNLONGIKRV 104
297 KEFLSSPNTQWELRAEYAVMHFSLTADRIDDILKVIYVDSMNHGARS 346
|||||
105 KEFLSSPNTQWELRAEYAVMHFSLTADRIDDILKVIYVDSMNHGARS 154
347 KLREELATLAEKLIYSYIOAEINKHLSSTGTINIHDSINIMDKNLXGY 396
|||||
155 KLREELATLAEKLIYSYIOAEINKHLSSTGTINIHDSINIMDKNLXGY 204
```

```
397 TDEEFKASAEYKILIEKMPOTTIOVDGSEKRIYSIKDFLGSERNRTGALG 446
|||||
205 TDEEFKASAEYKILIEKMPOTTIOVDGSEKRIYSIKDFLGSERNRTGALG 254
447 NLKNSYSYKNKNNELSHFATCSDKSRPLNDVSOQTOLSDISRFNSA 496
|||||
255 NLKNSYSYKNKNNELSHFATCSDKSRPLNDVSOQTOLSDISRFNSA 304
497 IEALNRFIOKYDSVMORLLDPTSGK 521
|||||
305 IEALNRFIOKYDSVMORLLDPTSGK 329

seq_name: N_Geneseq_36:T04223

seq_documentation_block:
ID T04223 standard; DNA; 1014 BP.
AC T04223:
DE 18-APR-1996 (first entry)
DE Partial lcrV (V antigen) gene of Y. pestis.
KW lcrV; V antigen; virulence; plaque; vaccine; epitope; ss.
OS Yersinia pestis.
FH Key location/Qualifiers
FT cds 1..990
FT /*tag= a
PN W09524475-A1.
PD 14-SEP-1995.
PF 06-MAR-1995: G00481.
PR 08-MAR-1994: GB-004577.
PA (MINA ) UK SEC FOR DEFENCE.
PI Leary SEC, Titball RW, Williamson ED, Leary SE;
DR WPI; 95-328268/42.
DR P-PSDB: R79962.
PT Recombinant DNA expressing Yersinia pestis V antigen - useful in
PT oral or parenteral vaccines for protection against plague
PS Claim 6: Page 15-16; 25pp; English.
CC T04222-23 are DNA sequences (lcrV) encoding all or a protective epitopic
CC part of the mature V protein of Yersinia pestis. The protein was
CC expressed as a fusion protein with maltose binding protein or
CC glutathione-S-transferase in 3 different plasmid vectors. Y. pestis
CC is the highly virulent causative organism of plague in a wide range of
CC animals, including man. The V antigen (lcrV) is an unstable 37.3 kDa
CC monomeric peptide encoded on the ca. 70 kb lcr plasmid. The V antigen
CC is postulated to act as a virulence antigen, and transformed
CC microorganisms contg. recombinant DNA encoding a V antigen protein/
CC peptide are useful in vaccines to protect against plague.
SQ Sequence 1014 BP; 343 A; 185 C; 204 G; 282 T;

alignment_scores:
Quality: 1441.00 Length: 325
Ratio: 4.434 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.385

alignment_block:
US-08-699-716A-2 x T04223 (1) ..
Align seg 1/1 to: T04223 from: 1 to: 1014

197 IRAYEONPOHFIEDELEKRVVEOLTGHGSSVLEELVOLVKDKNIDISTKYD 246
|||||
5 IRAYEONPOHFIEDELEKRVVEOLTGHGSSVLEELVOLVKDKNIDISTKYD 54
247 PRKQSEVFNANRVITDDIELLKKILAYFLPEPTILKGGHYDNLONGIKRV 296
|||||
55 PRKQSEVFNANRVITDDIELLKKILAYFLPEPTILKGGHYDNLONGIKRV 104
297 KEFLSSPNTQWELRAEYAVMHFSLTADRIDDILKVIYVDSMNHGARS 346
|||||
105 KEFLSSPNTQWELRAEYAVMHFSLTADRIDDILKVIYVDSMNHGARS 154
347 KLREELATLAEKLIYSYIOAEINKHLSSTGTINIHDSINIMDKNLXGY 396
|||||
155 KLREELATLAEKLIYSYIOAEINKHLSSTGTINIHDSINIMDKNLXGY 204
```

397 TDEEIFKAAEYKILEKMPQTIIQVDSKKIYSIDFJGSEKKRTGALG 446  
|||||  
205 TDEEIFKAAEYKILEKMPQTIIQVDSKKIYSIDFJGSEKKRTGALG 254  
447 NIKNSYSYKDKNNELSHFATCSDKSRPLNDLVSOQTQLOLSDTSFNRA 496  
|||||  
255 NIKNSYSYKDKNNELSHFATCSDKSRPLNDLVSOQTQLOLSDTSFNRA 304  
497 IEALNRFIOKYDSVMORLLDDTSGK 521  
|||||  
305 IEALNRFIOKYDSVMORLLDDTSGK 329

seq\_name: N\_Geneseq\_36:V41595

seq\_documentation\_block:  
ID V41595 standard: DNA; 510 BP.  
AC V41595:  
DE 26-OCT-1998 (first entry)  
DT Nucleotide sequence of F1 antigen nypfl(b)sec544.  
PS F1 antigen; plasmid; vaccine; plague; ds.  
OS Yersinia pestis.  
PN MO9824912-A2.  
PD 11-JUN-1998:  
PF 04-DEC-1997: U22617.  
PR 04-DEC-1996: US-767115.  
PA (HESK-) HESKA CORP.  
PI Haanes EJ, Osorio JE, Thomas RE;  
DR WPI: 98-333331/29.  
PT Protection of animals against plague - using nucleic acid encoding  
PT antigen from Yersinia, Pasteurella and Francisella spp.  
CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used  
CC in the method of the invention. Plasmid and host cells are used to  
CC produce recombinant antigens, especially Yersinia pestis antigens.  
CC The recombinant antigens can be used in vaccines that are capable of  
CC protecting an animal from contracting plague.  
SQ Sequence 510 BP; 153 A; 114 C; 106 G; 137 T;

alignment\_scores:  
Quality: 719.00 Length: 170  
Ratio: 4.229 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-08-699-716a-2 x V41595 (1) ..

Align seg 1/1 to: V41595 from: 1 to: 510

24 MKKISSVIAIAFGTATANAADLTASTATATLVEPARITLYKRGAPI 73  
|||||  
1 MKKISSVIAIAFGTATANAADLTASTATATLVEPARITLYKRGAPI 50  
74 TIMDNGNIDTELLVGTLLTGKGTGTTSTSVNFTDAAGDPMYLTFTSODG 123  
|||||  
51 TIMDNGNIDTELLVGTLLTGKGTGTTSTSVNFTDAAGDPMYLTFTSODG 100  
124 NNHOFITTKVIGKDSRDFDISPKVNGENLVGDVYVLTATGSDOFFVRSISGK 173  
|||||  
101 NNHOFITTKVIGKDSRDFDISPKVNGENLVGDVYVLTATGSDOFFVRSISGK 150  
174 GGLAAGKTYTDAVTVVSNQ 193  
|||||  
151 GGLAAGKTYTDAVTVVSNQ 170

seq\_name: N\_Geneseq\_36:Q92819

seq\_documentation\_block:  
ID Q92819 standard: DNA; 542 BP.  
AC Q92819:  
DT 17-DEC-1995 (first entry)  
DT Yersinia pestis cafl (F1) antigen in plasmid pFORF1b.

KW Vaccine; antigen; Salmonella typhimurium; Salmonella typhi;  
KW bubonic plague; pneumonic plague; ds.  
OS Yersinia pestis.  
FH Key  
FT cds  
FT Location/Qualifiers  
FT ..7  
FT /\*tag= a  
FT /note= "first protein encoded by pFORF1b"  
FT 1..6  
FT misc\_feature  
FT /\*tag= b  
FT 536..541  
FT misc\_feature  
FT /\*tag= c  
FT /note= "cafl open reading frame downstream seq."  
FT 21..530  
FT /\*tag= d  
FT /note= "cafl fusion protein"

MO9518231-A1.  
PD 06-JUL-1995.  
PF 23-DEC-1994: G02818.  
PR 24-DEC-1993: GB-026425.  
PA (MINA ) UK SEC FOR DEFENCE.  
PI Howells A, Leary SEC, Oyston PCF, Titball RW, Williamson ED;  
DR P-PSDB: R76528.  
PT DNA constructs capable of transforming microorganisms - which can be  
PT used as live or attenuated vaccines which induce an immune response,  
PT against Yersinia pestis, at mucosal surfaces.  
PS Disclosure: Page 19-20; 27pp: English.  
CC The sequence represents the plasmid pFORF1b including the entire  
CC Y. pestis cafl (F1) antigen gene having a 5' tail including a SacI  
CC restriction site, and up to TATAG downstream of the cafl ORF.  
CC The DNA construct can be used to transform human or animal gut  
CC colonizing microorganisms, specifically attenuated Salmonella  
CC typhimurium or Salmonella typhi. The transformed microorganisms  
CC can be used as live/attenuated vaccines which induce immune  
CC responses at mucosal surfaces. The vaccines provide protection  
CC against infection with Y. pestis, and are parenterally and orally  
CC active vaccines offering protection against bubonic and pneumonic  
CC plague.  
SQ Sequence 542 BP; 164 A; 118 C; 114 G; 146 T;

alignment\_scores:  
Quality: 719.00 Length: 170  
Ratio: 4.229 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-08-699-716a-2 x Q92819 (3) ..

Align seg 1/1 to: Q92819 from: 1 to: 542

24 MKKISSVIAIAFGTATANAADLTASTATATLVEPARITLYKRGAPI 73  
|||||  
7 MKKISSVIAIAFGTATANAADLTASTATATLVEPARITLYKRGAPI 56  
74 TIMDNGNIDTELLVGTLLTGKGTGTTSTSVNFTDAAGDPMYLTFTSODG 123  
|||||  
57 TIMDNGNIDTELLVGTLLTGKGTGTTSTSVNFTDAAGDPMYLTFTSODG 106  
124 NNHOFITTKVIGKDSRDFDISPKVNGENLVGDVYVLTATGSDOFFVRSISGK 173  
|||||  
107 NNHOFITTKVIGKDSRDFDISPKVNGENLVGDVYVLTATGSDOFFVRSISGK 156  
174 GGLAAGKTYTDAVTVVSNQ 193  
|||||  
157 GGLAAGKTYTDAVTVVSNQ 176

seq\_name: N\_Geneseq\_36:V41594

seq\_documentation\_block:  
ID V41594 standard: DNA; 544 BP.  
AC V41594:  
DT 26-OCT-1998 (first entry)

DE Nucleotide sequence of F1 antigen nypF1(a)sec544.  
KW F1 antigen; plasmid; vaccine; plague; ds.  
OS Yersinia pestis.  
FH Key Location/Qualifiers  
FT CDS 17..529  
FT /tag= a  
FT /product= "F1 antigen"  
PN WO9824912-A2.  
PD 11-JUN-1998.  
PF 04-DEC-1997; U22617.  
PR 04-DEC-1996; US-767115.  
PA (HESK-) HESKA CORP.  
PI Haanes EJ, Osorio JE, Thomas RE;  
DR WPI: 98-333331/29.  
DR P-PSDB: W59782.  
PT Protection of animals against plague - using nucleic acid encoding  
PT antigen from Yersinia, Pasteurella and Francisella spp.  
PS Claim 8; Page 51; 75pp; English.  
CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used  
CC in the method of the invention. Plasmid and host cells are used to  
CC produce recombinant antigens, especially Yersinia pestis antigens.  
CC The recombinant antigens can be used in vaccines that are capable of  
CC protecting an animal from contracting plague.  
SQ Sequence 544 BP; 164 A; 120 C; 114 G; 146 T;  
  
alignment\_scores:  
Quality: 719.00 Length: 170  
Ratio: 4.229 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
  
alignment\_block:  
US-08-699-716a-2 x V41594 (2) ..  
  
Align seg 1/1 to: V41594 from: 1 to: 544  
  
24 MKRISVIAIALFGTATANADLTASTATATLVEPARITLYKGCAP1 73  
|||||  
6 MKRISVIAIALFGTATANADLTASTATATLVEPARITLYKGCAP1 55  
|||||  
74 TIMDNCNIDTELLVGTLLGKYKGTSTSVNFTDAAGDPMYLFTFSQDG 123  
|||||  
56 TIMDNCNIDTELLVGTLLGKYKGTSTSVNFTDAAGDPMYLFTFSQDG 105  
|||||  
124 NNHQFTKVIKDSRDEDISPKVNGENLVGDVVLATGSDDFVRSIGSK 173  
|||||  
106 NNHQFTKVIKDSRDEDISPKVNGENLVGDVVLATGSDDFVRSIGSK 155  
|||||  
174 GCKLAGKXTDAVTYVSNO 193  
|||||  
156 GCKLAGKXTDAVTYVSNO 175  
  
seq\_name: N\_Geneseq\_36:V41596  
  
seq\_documentation\_block:  
ID V41596 standard: DNA: 544 BP.  
AC V41596.  
DT 26-OCT-1998 (first entry)  
DE Nucleotide sequence of F1 antigen nypF1sec510.  
KW F1 antigen; plasmid; vaccine; plague; ds.  
OS Yersinia pestis.  
FH Key Location/Qualifiers  
FT CDS 17..532  
FT /tag= a  
FT /product= "F1 antigen"  
PN WO9824912-A2.  
PD 11-JUN-1998.  
PF 04-DEC-1997; U22617.  
PR 04-DEC-1996; US-767115.  
PA (HESK-) HESKA CORP.  
PI Haanes EJ, Osorio JE, Thomas RE;  
DR WPI: 98-333331/29.  
DR P-PSDB: W59783.

PT Protection of animals against plague - using nucleic acid encoding  
PT antigen from Yersinia, Pasteurella and Francisella spp.  
PS Claim 8; Pages 53-54; 75pp; English.  
CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used  
CC in the method of the invention. Plasmid and host cells are used to  
CC produce recombinant antigens, especially Yersinia pestis antigens.  
CC The recombinant antigens can be used in vaccines that are capable of  
CC protecting an animal from contracting plague.  
SQ Sequence 544 BP; 166 A; 118 C; 112 G; 148 T;  
  
alignment\_scores:  
Quality: 719.00 Length: 170  
Ratio: 4.229 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
  
alignment\_block:  
US-08-699-716a-2 x V41596 (2) ..  
  
Align seg 1/1 to: V41596 from: 1 to: 544  
  
24 MKRISVIAIALFGTATANADLTASTATATLVEPARITLYKGCAP1 73  
|||||  
6 MKRISVIAIALFGTATANADLTASTATATLVEPARITLYKGCAP1 55  
|||||  
74 TIMDNCNIDTELLVGTLLGKYKGTSTSVNFTDAAGDPMYLFTFSQDG 123  
|||||  
56 TIMDNCNIDTELLVGTLLGKYKGTSTSVNFTDAAGDPMYLFTFSQDG 105  
|||||  
124 NNHQFTKVIKDSRDEDISPKVNGENLVGDVVLATGSDDFVRSIGSK 173  
|||||  
106 NNHQFTKVIKDSRDEDISPKVNGENLVGDVVLATGSDDFVRSIGSK 155  
|||||  
174 GCKLAGKXTDAVTYVSNO 193  
|||||  
156 GCKLAGKXTDAVTYVSNO 175  
  
seq\_name: N\_Geneseq\_36:T38248  
  
seq\_documentation\_block:  
ID T38248 standard: DNA: 547 BP.  
AC T38248.  
DT 28-DEC-1996 (first entry)  
DE Y. pestis F1 antigen caf1 gene (including signal sequence).  
KW Plague; vaccine; genetic immunisation; F1 antigen; caf1;  
KW V antigen; ds.  
OS Yersinia pestis strain GB.  
FH Key Location/Qualifiers  
FT Signal\_peptide 29..91  
FT /tag= a  
FT mat\_peptide 92..538  
FT /tag= b  
PN WO9628551-A1.  
PD 19-SEP-1996.  
PF 13-MAR-1996; G00571.  
PR 13-MAR-1995; GB-005059.  
PR 15-SEP-1995; GB-018946.  
PR 05-DEC-1995; GB-024825.  
PA (MINA ) UK SEC FOR DEFENCE.  
PI Bennett AM, Leary SEC, Oyston PCF, Titball RW, Williamson ED;  
DR WPI: 96-433824/43.  
DR P-PSDB: W01043.  
PT Yersinia pestis V antigen and F1 antigen or their protective  
PT epitopic parts - useful in vaccine for protection against plague  
PS Claim 41; Page 61-62; 98pp; English.  
CC A DNA sequence (T38248) comprises the caf1 gene, including the  
CC signal sequence, coding for the F1 antigen (W01043) of Yersinia  
CC pestis. It was obtd. by PCR amplification (see also T38257-58)  
CC of Y. pestis DNA. The PCR product was cloned into plasmid pBKCMV  
CC and the resulting plasmid (pT1A8) was used to transform E. coli Nova  
CC Blue. Purified plasmid, when administered by i.m. injection,  
CC induced an immunoglobulin response to F1 in Balb/c mice. Live  
CC vaccines comprising gut colonising organisms transformed with the



```
FT      /product= "F1 antigen"
PN      W09824912-A2.
PD      11-JUN-1998.
PF      04-DEC-1997: U22617.
PR      04-DEC-1996: US-767115.
PA      (HESK-) HESKA CORP
PI      Hanes EJ, Osorio JE, Thomas RE;
DR      WPI; 98-33331/29.
DR      P-PSDB; W59787.
PT      Protection of animals against plague - using nucleic acid encoding
CC      antigen from Yersinia, Pasteurella and Francisella spp.
CC      Claim 8, Pages 59-60; 75pp; English.
CC      This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
CC      in the method of the invention. Plasmid and host cells are used to
CC      produce recombinant antigens, especially Yersinia pestis antigens.
CC      The recombinant antigens can be used in vaccines that are capable of
CC      protecting an animal from contracting plague.
SQ      Sequence 474 BP; 143 A; 106 C; 99 G; 126 T;

alignment_scores:
      Quality: 634.00      Length: 149
      Ratio: 4.255      Gaps: 0
      Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-08-699-716A-2 x V41600 (1)  ..
Align seg 1/1 to: V41600 from: 1 to: 474

45 ADLTAATATATATLVEPARITLTYKKGAPITIMDNGNIDITELLVGTLLG 94
   |||||||
4 ADLTAATATATATLVEPARITLTYKKGAPITIMDNGNIDITELLVGTLLG 53
95 YKTGTTSTSVNFTDAAGDPMYLFTFSQDGNHOFKTKVIGKSRDPDISP 144
   |||||||
54 YKTGTTSTSVNFTDAAGDPMYLFTFSQDGNHOFKTKVIGKSRDPDISP 103
145 KVGNGENLVGDDVVLATGSDOFFVRSIGSKGKLAGKYTDAVTIVYSNQ 193
   |||||||
104 KVGNGENLVGDDVVLATGSDOFFVRSIGSKGKLAGKYTDAVTIVYSNQ 152

seq_name: N_Geneseq_36:Q92817

seq_documentation_block:
ID      Q92817 standard: DNA; 541 BP.
AC      Q92817;
DT      17-DEC-1995 (first entry)
DE      Yersinia pestis caf1 (F1) antigen in plasmid pFGAL2a.
KW      vaccine; antigen; Salmonella typhimurium; Salmonella typhi;
      KW bubonic plague; pneumonic plague; ds.
OS      Yersinia pestis.
FH      Key
FT      cds      2..454      location/Qualifiers
FT      misc_feature 1..6      /*tag= a
FT      misc_feature /*tag= b
FT      misc_feature /*note= "lacZ promoter fusion site"
FT      misc_feature 536..541
FT      misc_feature /*tag= C
FT      misc_feature /*note= "vector pFGAL2a bases"
PN      W09518231-A1.
PD      06-JUL-1995.
PF      23-DEC-1994: G02818.
PR      24-DEC-1993: GB-026425.
PA      (MINA ) UK SEC FOR DEFENCE.
PI      Howells A, Leaty SEC, Oyston PCF, Tlball RW, Williamson ED;
DR      WPI; 95-246396/32.
DR      P-PSDB; R76526.
PT      DNA constructs capable of transforming microorganisms - which can be
CC      used as live or attenuated vaccines which induce an immune response,
CC      against Yersinia pestis, at mucosal surfaces.
PS      Claim 7; Page 15-16; 27pp; English.
```

```
CC      The sequence represents the plasmid pFGAL2a construct showing the
CC      fusion of the first few bases of beta-galactosidase in the vector
CC      with the Y. pestis caf1 (F1) antigen minus its signal sequence and
CC      having a 5' tail including a SacI restriction site, and up to the
CC      caf1 AAC-3' end with some vector bases. The DNA construct can be
CC      used to transform human or animal gut colonizing microorganisms,
CC      specifically attenuated Salmonella typhimurium or Salmonella typhi.
CC      The transformed microorganisms can be used as live/attenuated
CC      vaccines which induce immune responses at mucosal surfaces. The
CC      vaccines provide protection against infection with Y. pestis, and
CC      are parenterally and orally active vaccines offering protection
CC      against bubonic and pneumonic plague.
SQ      Sequence 541 BP; 163 A; 120 C; 111 G; 147 T;

alignment_scores:
      Quality: 632.00      Length: 151
      Ratio: 4.185      Gaps: 0
      Percent Similarity: 100.000      Percent Identity: 98.013

alignment_block:
US-08-699-716A-2 x Q92817 (2)  ..
Align seg 1/1 to: Q92817 from: 1 to: 541

43 NAADLTAATATATLVEPARITLTYKKGAPITIMDNGNIDITELLVGTLL 92
   ::|||||
1 SSADLTAATATATLVEPARITLTYKKGAPITIMDNGNIDITELLVGTLL 50
93 GGYKGTGTTSTSVNFTDAAGDPMYLFTFSQDGNHOFKTKVIGKSRDPDI 142
   |||||||
51 GGYKGTGTTSTSVNFTDAAGDPMYLFTFSQDGNHOFKTKVIGKSRDPDI 100
143 SPKVGNGENLVGDDVVLATGSDOFFVRSIGSKGKLAGKYTDAVTIVYSN 192
   |||||||
101 SPKVGNGENLVGDDVVLATGSDOFFVRSIGSKGKLAGKYTDAVTIVYSN 150
193 Q 193
151 Q 151
```